

**IDENTIFICATION AND
ANALYSIS OF THE
FLAGELLIN GENE AND PROTEIN
FROM THE GENUS
*PECTINATUS***

A Thesis Submitted to the College of
Graduate Studies and Research
in Partial Fulfillment of the Requirements
for the Degree of Master of Science
in the Department of Microbiology and Immunology
University of Saskatchewan
Saskatoon

By
Bonnie Chaban

PERMISSION TO USE

In presenting this thesis in partial fulfillment of the requirements for a Postgraduate degree from the University of Saskatchewan, I agree that the Libraries of the University of Saskatchewan may make it freely available for inspection. I further agree that permission for copying of this thesis in any manner, in whole or in part, for scholarly purposes may be granted by the professor or professors who supervised my thesis work or, in their absence, by the Head of the Department of Microbiology and Immunology or the Dean of the College of Graduate Studies and Research in which my thesis work was done. It is understood that any copying or publication or use of this thesis or parts thereof for financial gain shall not be allowed without my written permission. It is also understood that due recognition shall be given to me and to the University of Saskatchewan in any scholarly use which may be made of any material in my thesis.

Requests for permission to copy or to make other use of material in this thesis in whole or part should be addressed to:

The Head of the Department of Microbiology and Immunology

University of Saskatchewan

Saskatoon, Saskatchewan

S7N 5E5

ABSTRACT

The use of reduced oxygen-packaging techniques has resulted in anaerobic bacteria emerging as a problem for the brewing industry over the last twenty-five years. The genus *Pectinatus*, consisting of the species *P. cerevisiiphilus* and *P. frisingensis*, is a concern for producers of unpasteurized beer. As a result, there is an ongoing need to both understand this genus and develop rapid detection methodologies to combat its presence in the brewery. The objectives of this study were to sequence and characterize the flagellin genes from both *Pectinatus* species and evaluate the genes and proteins from a taxonomic and detection-suitability standpoint.

A combination of micro-protein sequencing, polymerase chain reaction (PCR) and Bubble-PCR was used to completely sequence one flagellin gene from each *Pectinatus* species. This knowledge was then utilized to sequence the flagellin gene from four additional *Pectinatus* isolates, two from each species. To confirm the identity of the flagellin genes, one flagellin gene from each species was cloned, expressed and detected with *Pectinatus*-specific antibodies. A discrepancy between of the predicted protein size and the actual protein size led to tests for glycosylation, a post-translational modification. Taxonomic analyses, based on the flagellin genes, were conducted at both the superkingdom and genus levels. Finally, genus- and species-specific PCR primer sets were designed and tested for the specific detection of *Pectinatus* in the brewery.

Cloning and expression data confirmed the identity of the sequenced genes as *Pectinatus* flagellin genes. Glycosylation was positively confirmed to be a post-translational modification for five of the six strains tested. Phylogenetic analysis revealed that both of the *Pectinatus* species grouped with the phylum *Firmicutes* (low G+C, Gram-positive bacteria) and that there was more diversity at the species level

within the *P. frisingensis* flagellin gene than the *P. cerevisiophilus* flagellin gene. As a final point, the detection of most *Pectinatus* isolates was achieved with the preliminary PCR primer sets designed, however, some non-*Pectinatus* beer spoilage organisms, primarily wort spoilage organisms, were also detected. Both the basic science and the applied results generated from this study will aid the brewing industry in its ongoing battle to control *Pectinatus* contamination.

ACKNOWLEDGEMENTS

I have lots of people to thank for all the guidance and support I've received throughout this M.Sc. project. To begin, I thank Dr. Barry Ziola, my supervisor, for his unending patience and unwavering support. As well, my committee, consisting of Drs. Harry Deneer, Sean Hemmingsen and Wei Xiao, has been an excellent resource and I appreciate all the assistance each member has given me.

My thanks also go out to all the Department of Microbiology and Immunology faculty and support staff. Everyone has been a part of my experience here and the assistance I've received and friendships I've made will never be forgotten. Along the same lines, I owe a huge thank you to all the graduate students in the department for their suggestions, discussions and friendships. Among other things, they've helped me to never underestimate the number of problems that can be solved at Louis.

I also acknowledge the DNA and protein sequencing work done at the Plant Biotechnology Institute and University of Victoria, respectively. As well, I appreciate all the contributions made by other members of the BZ lab during my time here.

I thank NSERC (National Sciences and Engineering Research Council) and the College of Medicine for scholarships awarded to me and Coors Brewing Company for supporting this research. As well, I thank CIHR (Canadian Institute for Health Research), the Burroughs Wellcome Fund, the University of Saskatchewan, the College of Medicine, the Department of Microbiology and Immunology and Dr. Ziola for financial support to attend Canadian Bioinformatics Workshops.

This thesis is dedicated in memory of my mother, Laura Chaban. So much of who I am is due to her. I hope she'd be proud.

TABLE OF CONTENTS

PERMISSION TO USE	i
ABSTRACT	ii
ACKNOWLEDGEMENTS	iv
TABLE OF CONTENTS	v
LIST OF TABLES	viii
LIST OF FIGURES	ix
LIST OF ABBREVIATIONS	x
 1.0 LITERATURE REVIEW / INTRODUCTION	 1
1.1 Beer spoilage bacteria	1
1.2 <i>Pectinatus</i>	1
1.2.1 Basic characteristics	2
1.2.2 Taxonomy of <i>Pectinatus</i>	5
1.2.3 Methods for detection of <i>Pectinatus</i>	6
1.2.3.1 Detection of whole organism or by-products	6
1.2.3.2 Detection by DNA-based methods	8
1.3 Flagella	9
1.4 Flagellin	10
1.4.1 Basic characteristics	10
1.4.2 Regulatory elements	13
1.4.3 Post-translational modifications	15
1.4.3.1 Glycosylation	16
1.4.3.2 Phosphorylation	17
1.4.3.3 Sulfation	17
1.4.3.4 Flagella sheaths	17
1.5 Research objectives	18
 2.0 MATERIALS AND METHODS	 20
2.1 Propagation of bacteria	20
2.2 Isolation of flagellin protein	24
2.3 Preparation of flagellin protein for N-terminal sequencing	24
2.4 Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE)	25
2.5 N-terminal amino acid micro-sequencing	25
2.6 DNA extraction	26
2.7 Polymerase chain reaction (PCR)	26
2.7.1 Reaction components	26
2.7.2 Primer design	27
2.7.3 PCR programs	31
2.7.4 Bubble-PCR	32
2.8 PCR product visualization and purification	34
2.9 DNA sequencing	36
2.10 Sequence analysis and assembly	36
2.11 Cloning of the flagellin genes	37
2.11.1 Glutathione S-transferase (GST) gene fusion system	37
2.11.2 Cloning procedure	37

2.12 Detection of GST-flagellin fusion proteins.....	41
2.12.1 Extraction of fusion proteins.....	41
2.12.2 Immunoblotting of fusion proteins	41
2.13 Carbohydrate detection	43
2.13.1 Carbohydrate-specific stains	43
2.13.2 Glycosidases.....	44
2.14 Phylogenetic analysis.....	45
3.0 RESULTS	48
3.1 Flagellin protein micro-sequencing.....	48
3.2 Initial sequence of flagellin genes.....	48
3.2.1 <i>Pectinatus cerevisiophilus</i>	48
3.2.2 <i>Pectinatus frisingensis</i>	53
3.3 Expression and confirmation of flagellin genes.....	53
3.4 <i>Pectinatus</i> flagellin characteristics.....	57
3.5 Glycosylation analysis via staining and glycosidases	62
3.6 Taxonomic analysis of <i>Pectinatus</i> via phylogenetics	67
3.7 Genus- and Species-Specific PCR diagnostics	67
4.0 DISCUSSION	74
4.1 <i>Pectinatus</i> flagellin.....	74
4.1.1 Sequencing the flagellin.....	74
4.1.2 Characteristics of the flagellin open reading frame	77
4.1.3 Regulatory elements.....	79
4.1.4 Glycosylation	81
4.2 Phylogeny.....	85
4.2.1 Bacteria (superkingdom) level	87
4.2.2 <i>Pectinatus</i> (genus) level	88
4.3 Detection of <i>Pectinatus</i> via the flagellin gene	89
4.4 Conclusions	91
4.5 Future considerations	94
5.0 BIBLIOGRAPHY	97
6.0 APPENDIX A - Solutions, Buffers and Media.....	106
7.0 APPENDIX B	109
7.1 Chemical, media and reagents.....	109
7.2 Equipment and supplies	113
7.3 Software used	115
8.0 APPENDIX C.....	CD
8.1 Unaligned sequences (in FASTA format)	
8.1.1 All flagellin proteins	
8.1.2 <i>Pectinatus</i> flagellin DNA	
8.1.3 <i>Pectinatus</i> flagellin protein	
8.1.4 <i>Pectinatus</i> 16S sequences	

8.2 Aligned sequences (in GeneDoc format)

8.2.1 All flagellin proteins

8.2.2 *Pectinatus* flagellin DNA

8.2.3 *Pectinatus* flagellin protein

8.2.4 *Pectinatus* 16S rRNA

8.3 Distance matrices (in plain text format)

8.3.1 All flagellin proteins

8.3.2 *Pectinatus* flagellin DNA

8.3.3 *Pectinatus* flagellin protein

8.3.4 *Pectinatus* 16S rRNA

LIST OF TABLES

Table 1 – <i>Pectinatus</i> isolates	21
Table 2 – Flagellin gene and protein sequences obtained from GenBank	22
Table 3 – Beer spoilage isolates used for genus- and species-specific PCR analysis.....	23
Table 4 – PCR primers	28
Table 5 – PCR product composition of <i>Pectinatus cerevisiiphilus</i> ATCC 29359 ^T	51
Table 6 – PCR product composition of <i>Pectinatus frisingensis</i> DSM 20465	54
Table 7 – Summary of carbohydrate detection	64
Table 8 – Summary of genus- and species-specific PCR results	73

LIST OF FIGURES

Figure 1 – Generic structure of a flagellin gene.....	12
Figure 2 – Diagram of Bubble-PCR methodology	33
Figure 3 - Map of the glutathione S-transferase fusion vector, pGEX-6P-1, showing the reading frame and main features.....	38
Figure 4 - Alpha-chymotrypsin digestion of <i>Pectinatus</i> flagellin protein	49
Figure 5 – Schematic of the PCR assembly of <i>Pectinatus cerevisiiphilus</i> ATCC 29359 ^T flagellin gene.....	52
Figure 6 – Schematic of the PCR assembly of <i>Pectinatus frisingensis</i> DSM 20466 flagellin gene.....	55
Figure 7 – Immunoblots of GST-flagellin fusion proteins.....	56
Figure 8 – ClustalX alignment of <i>Pectinatus</i> flagellin gene sequences.....	58
Figure 9 – ClustalX alignment of <i>Pectinatus</i> flagellin proteins.....	61
Figure 10 – Sigma 28 flagellin promoters	63
Figure 11 – Carbohydrate staining results	65
Figure 12 – Phylogenetic tree of flagellin protein sequences across the bacterial superkingdom.....	68
Figure 13 – Phylogenetic trees of <i>Pectinatus</i> flagellin at the DNA and protein level.....	69
Figure 14 – Phylogenetic tree of <i>Pectinatus</i> 16S rRNA sequences	70
Figure 15 – Binding locations for PCR primers Pect550, Pf253 and Pc185	72

LIST OF ABBREVIATIONS

aa	Amino acids
amp	Ampicillin
Asn	Asparagine
ATCC	American Type Culture Collection
<i>B. burgdorferi</i>	<i>Borrelia burgdorferi</i>
BLAST	Basic local alignment search tool
bp	Base pair(s)
BSA	Bovine serum albumin
<i>B. subtilis</i>	<i>Bacillus subtilis</i>
COG	Clusters of Orthologous Groups of proteins database
DEFT	Direct epifluorescent staining technique
dNTPs	Deoxyribonucleic triphosphates
DSM	German Collection of Microorganisms and Cell Cultures
<i>E. coli</i>	<i>Escherichia coli</i>
EDTA	Ethylenediaminetetraacetic acid
EMBOSS	European Molecular Biology Open Software Suite
Fet	Fetuin
GlcNAc	N-acetyl- β -glucosamine
GST	Glutathione S-transferase
HRP	Horse radish peroxidase
Ig	Immunoglobulin
LB	Luria-Bertani
LPS	Lipopolysaccharide

mRNA	Messenger RNA
MRS	De Man, Rogosa, Sharpe
Ova	Ovalbumin
<i>P. aeruginosa</i>	<i>Pseudomonas aeruginosa</i>
PAM	Point-accepted mutation
PAS	Periodic Acid-Schiff's stain
PBS	Phosphate-buffered saline
Pc	<i>Pectinatus cerevisiiphilus</i>
<i>P. cerevisiiphilus</i>	<i>Pectinatus cerevisiiphilus</i>
PCR	Polymerase chain reaction
Pf	<i>Pectinatus frisingensis</i>
Pfam	Protein families database
<i>P. frisingensis</i>	<i>Pectinatus frisingensis</i>
PHYLP	Phylogeny Inference Package
PNGase F	Peptide: N-Glycosidase F
PVDF	Polyvinylidene difluoride
<i>R. cecicola</i>	<i>Roseburia cecicola</i>
rdH ₂ O	Reverse osmosed, deionized water
rDNA	Ribosomal DNA
rRNA	Ribosomal RNA
SDS-PAGE	Sodium dodecyl sulfate-polyacrylamide gel electrophoresis
Ser	Serine
SSC	Saline-sodium citrate

<i>S. typhimurium</i>	<i>Salmonella typhimurium</i>
TBE	Tris-Boric acid-EDTA buffer
TBS	Tris-buffered saline
TE	Tris-EDTA
Thr	Threonine
Tris	Tris(hydroxymethyl)aminomethane
TTBS	TBS, containing 0.05% (v/v) Tween 20
UV	Ultraviolet
VTT	Technical Research Centre of Finland

1.0 LITERATURE REVIEW / INTRODUCTION

1.1 Beer spoilage bacteria

Human beings have been making and enjoying alcoholic beverages for thousands of years. Today, brewing is a multi-billion dollar industry. As with any other food or beverage, spoilage of beer by microorganisms is an ongoing concern. Only a select range of organisms can survive the low pH, alcohol content, hop residuals and lack of oxygen in packaged beer (Lee *et al.*, 1978). These “beer-spoilage bacteria” fall into five main categories: wort spoilers, lactic acid bacteria, acetic acid bacteria, *Zymomonas* and anaerobic Gram-negative bacteria. While the first four categories have been recognized, in one form or another, for a century or more, anaerobic Gram-negative bacteria were regarded as contaminants only in the early part of the fermentation process and not considered important beer spoilage organisms (Schleifer *et al.*, 1990). Because of this, they were generally ignored. This opinion began to change 25 years ago with the identification of a genuine anaerobic Gram-negative beer-spoilage organism, namely *Pectinatus* (Lee *et al.*, 1978). Since then, the category of anaerobic beer spoilers has grown to include the genera *Megasphaera*, *Selenomonas* and *Zymophilus* (Priest, 1996).

1.2 *Pectinatus*

Improved bottling techniques have reduced the oxygen levels inside packaged beer. This is desirable because the chemical breakdown of beer is slowed under oxygen-

reduced conditions, thus extending the shelf-life of packaged beer. Unfortunately, these bottling techniques have also allowed anaerobic Gram-negative bacteria to emerge as brewing spoilage organisms. The first of these new anaerobes was identified in the USA, in 1978, as a Gram-negative rod which did not fall into any known genus. It was designated as *Pectinatus cerevisiiphilus* (Lee *et al.*, 1978). The name, translated from the Latin, means combed (from its flagella arrangement), beer-loving bacteria. Since then, *Pectinatus* has caused spoilage in many other countries, including Finland, Germany, Japan, Norway and Sweden (Haikara, 1992). This bacterium produces a range of organic acids and sulfur-containing compounds that cause turbidity, bad odours and off-flavors. In 1990, the genus was expanded to include a second species, namely, *Pectinatus frisingensis* (Schleifer *et al.*, 1990). The new species name, *frisingensis*, is the Latin name of Freising, the town in Germany where the organism was initially isolated.

1.2.1 Basic characteristics

P. cerevisiiphilus and *P. frisingensis* are Gram-negative, motile, slightly curved rods, 0.7-0.8 μm in diameter and 2-32 μm in length. They occur singly, in pairs and very rarely in short chains. Long, helical filaments are characteristic of older cells. They are obligate anaerobes, non-spore forming and have an optimum growth temperature of 30⁰C. *Pectinatus* isolates produce acetic, lactic, propionic and succinic acids, as well as hydrogen sulphide, methyl mercaptane and acetoin. The mol % G+C content ranges from 38 to 41 (Lee *et al.*, 1978, 1981; Haikara *et al.*, 1981a; Schleifer *et al.*, 1990). Fatty acid analysis shows a distribution dominated by odd-numbered chains,

with temperature playing a large role in the saturated to unsaturated fatty acid ratio (Helander and Haikara, 1995; Flahaut *et al.*, 2000). The most striking visual characteristic of the genus is the comb-like arrangement of its flagella along the concave side of the bacteria (Lee *et al.*, 1978; Schleifer *et al.*, 1990). As well, *Pectinatus* species possess a very thick peptidoglycan layer and have cytoplasmic membrane characteristics that are typical of Gram-positive bacteria. Conversely, they possess an outer membrane typical of Gram-negative bacteria and stain Gram-negative to Gram-variable (Lee *et al.*, 1978). These properties make *Pectinatus* an intermediate form between the two classically distinguished groups of bacteria (Haikara *et al.*, 1992).

Since *Pectinatus* species are beer-spoilage organisms, they must tolerate the conditions found in beer. *Pectinatus* is, therefore, very acid-tolerant, with pH values less than 4.1 required to see any retardation of growth (Haikara, 1992). This acid-resistant characteristic has been linked to the structure of the lipid A-polysaccharide linkage in *Pectinatus*, which is of the asymmetric hexaacyl type (Helander *et al.*, 1994). *Pectinatus* grows quite well in beer with alcohol contents from 3.7-4.4% (w/v) and it is not until beers exceed an alcohol content of 5.2% (w/v) that growth is inhibited (Haikara, 1992). Both species in the genus are resistance to hop substances and have strict oxygen-free requirements (Haikara, 1992).

There are a number of phenotypic characteristics that differentiate the two *Pectinatus* species. Fermentation patterns reveal differences found in the degradation of xylose, cellobiose, melibiose and *N*-acetylglucosamine (Schleifer *et al.*, 1990). Growth studies show that *P. frisingensis* has a higher alcohol resistance, is more heat tolerant and has a wider pH tolerance range than *P. cerevisiiphilus*, suggesting that of the two *Pectinatus* species, *P. frisingensis* would be more likely to grow in beer (Tholozan *et al.*,

1997; Flahaut *et al.*, 2000). Lipopolysaccharide (LPS) analysis also illustrates differences. *P. cerevisiophilus* LPS is made up of a disaccharide consisting of D-glucose and D-fucose (in the furanose form). *P. frisingensis* LPS is a homopolymer consisting of 6-deoxy-L-altrofuranose (Senchenkova *et al.*, 1995). It is worth noting that D-fucose is a rare sugar and that this is the first recorded occurrence of 6-deoxy-L-altrofuranose in nature. Finally, striking differences in immunological characteristics exist between species (and within species), allowing for species differentiation to be based on surface antibody reactivity (Haikara *et al.*, 1981b; Gares *et al.*, 1993).

The natural habitat of *Pectinatus* is unknown. A possible clue to its natural environment is the presence of D-fucose in its LPS layer. This relatively rare sugar has traditionally been associated with plant pathogens. This suggests that *Pectinatus* may naturally occur as a plant pathogen and might have been carried into the brewery with raw ingredients such as malt or hops (Helander *et al.*, 1992). Since entering the brewery setting, *Pectinatus* has been identified in lubrication oil mixed with beer, chain lubricants, in drainage and water pipe systems, on the floor and ceiling of the filling hall, in the air of the filling hall, in the filling machine and in steeping water of malt before milling (Lee *et al.*, 1980; Haikara, 1992 and references within).

Although isolates appear to be ubiquitous throughout the brewery, *Pectinatus* is an easy organism to kill. Cells will not survive pasteurization and treatments of 58-60°C for one minute are sufficient to kill all *Pectinatus* cells (Lee *et al.*, 1980). For unpasteurized beer, proper cleaning and disinfection of brewery equipment is paramount. Lee *et al.* (1980) reports iodine and chlorine as effective disinfectants, whereas Haikara (1992) claims iodine to be ineffective and proposes peracetic acid (0.75%, 1 min) or formaldehyde (0.3%, 1 min) as suitable alternatives.

1.2.2 Taxonomy of *Pectinatus*

It is a daunting task to find an appropriate location for a new genus within the vast taxonomy of microorganisms. When the genus *Pectinatus* was first established, it resembled the organisms from two different families; *Bacillaceae* (under the Gram-positive phylum *Firmicutes*) and *Bacteroidaceae* (under the green sulfur bacteria phylum *Bacteroidetes/Chlorobi*). Since *Pectinatus* did not produce spores, which was considered a major characteristic within *Bacillaceae*, the new genus was placed under *Bacteroidaceae* (Lee *et al.*, 1978).

However, evidence accumulated indicating that *Pectinatus* belonged under the phylum *Firmicutes*. In 1981(b), Haikara *et al.* demonstrated that *Pectinatus* species synthesize propionic acid through a different pathway than *Bacteriodes* species (within *Bacteroidaceae*). Next, a phylogenetic analysis based on the 16S ribosomal RNA (rRNA) region revealed that *Pectinatus* is more closely related to *Selenomonas* and *Bacillus* species (within *Firmicutes*) than to anything from the *Bacteroidaceae* family (Schleifer *et al.*, 1990). About the same time, other members from *Bacteroidaceae* were being moved to the phylum *Firmicutes* based on 16S rRNA studies (Willems and Collins, 1995). Structural studies of the LPS within the genus *Pectinatus* were also demonstrating that both the fatty acid composition and the inner-core region composition of LPS are distinctly different between *Bacteriodes* and *Pectinatus* (Helander *et al.*, 1983, 1993). The phylogenetic similarity of *Pectinatus* to organisms located in the phylum *Firmicutes* continued with a relationship to *Selenomonas* and *Zymophilus* demonstrated through the 16S-23S ribosomal DNA (rDNA) interspacer region (Motoyama and Ogata, 2000a). Finally, the latest edition of Bergey's Manual of Systematic Bacteriology (Garrrity and Holt, 2001) officially moved *Pectinatus* into the

family *Acidaminococcaceae* (under the phylum *Firmicutes*) with *Megasphaera*, *Selenomonas* and *Zymophilus*.

Also in relation to taxonomy, it has been suggested that more than two species exist within the genus *Pectinatus*. Ribotyping studies, conducted with the largest collection of *Pectinatus* isolates to date, suggest that a third species within the genus *Pectinatus* is warranted from isolates conventionally identified as *P. cerevisiiphilus* (Suihko and Haikara, 2001). Continued research is needed to support this possibility.

1.2.3 Methods for detection of *Pectinatus*

Detection and identification have been the driving forces behind research into this beer-spoilage genus. Since *Pectinatus* emerged as a problem for brewers, a variety of techniques have been used to identify it as quickly as possible. This is understandable, since it took 30 days of incubation to isolate the original strain of *Pectinatus* (Lee *et al.*, 1978)! The approaches to detect *Pectinatus* isolates can be divided into two types of methodologies; namely, methods for detecting the whole organism or its by-products and methods for detecting microbial DNA.

1.2.3.1 Detection of whole organism or by-products

A number of classical methodologies have been used or proposed to detect beer-spoilage organisms. These methods are reviewed and compared by Day (1987) and Barney and Kot (1992). Included are direct methods, such as direct growth counts and staining, and indirect methods, such as pH change of the media, ATP bioluminescence, conductance and direct epifluorescent staining technique (DEFT). Two major problems exist when trying to apply these techniques to the specific detection and identification of

Pectinatus isolates. The first is the slow growth rate and anaerobic nature of *Pectinatus*. The second is inherent lack of organism specificity of many of the techniques. The direct methods, which are traditionally used in the brewery setting, rely on the growth of the organism. For *Pectinatus* isolates to be detected and identified in this manner, the brewery would need the apparatus and knowledge to grow anaerobic organisms and a growth time-frame of 2-3 weeks. Indirect methods offer little advantage, since pH change of the media, ATP bioluminescence and conductance all measure properties of actively growing cultures (by-product secretion, ATP utilization and conversion of uncharged substrate molecules to charged end-product molecules in the medium, respectively). These methods offer little utility for the detection of slow-growing cultures and no utility at all in the identification of the spoilage organism in question.

A number of the classical detection methodologies have been tested or adapted to detect and identify *Pectinatus*. DEFT, with acridine orange, was the first of these (Haikara, 1985). The method couples “forced beer” incubation (enrichment of beer with a selective media for *Pectinatus* and *Megasphaera*) with traditional DEFT detection. Forced beer is filtered through a membrane to concentrate any organisms present and the membrane is stained to detect cells. This method allows for detection of *Pectinatus* contaminates one to four days before visual turbidity of the beer, but could not guarantee specificity for *Pectinatus*, since all microorganisms present are stained by the technique. The detection limit of this method is approximately 1000 cells per filter, thus making the pre-enrichment step the dependent factor.

To have identification, as well as detection, membrane filter-based fluoroimmunoassay with surface-reactive monoclonal antibodies was put forward (Gares *et al.*, 1993). All microorganisms from a volume of beer are concentrated onto a

membrane by filtration. *Pectinatus* isolates are then specifically detected by fluoroimmunoassay using monoclonal antibodies. This allows the method to reach a detection limit of 2-4 cells in 10 ml of beer in less than 3 hours and positively identify *Pectinatus* organisms. The drawback to deploying this method is the initial difficulty in raising and testing the monoclonal antibodies needed to give identification in the assay.

Improvements to the growth medium for *Pectinatus* isolates have also been made. A selective medium to grow *Megasphaera* and *Pectinatus* isolates (called SMMP) was introduced in 1994 (Lee, 1994). In SMMP medium, *Megasphaera* and *Pectinatus* are selected for and grow within 3-4 days. This offers breweries without more sophisticated equipment the ability to grow (though not positively identify) *Pectinatus* in a more reasonable time-frame.

Gas chromatography is as another possible method for positive identification of *Pectinatus* isolates. The production of large quantities of propionic and acetic acid are unique characteristics of actively growing *Pectinatus* cells and these acids can be detected by gas chromatography analysis (Foster and Andersen, 1999). Although this method offers specificity in identification, the organism must be grown before analysis can be done and highly specialized equipment (gas chromatography apparatus) is needed.

1.2.3.2 Detection by DNA-based methods

Starting in the late 1990's, DNA-based molecular methodologies were introduced as detection strategies throughout the brewery quality control process. This is most likely due to advances in molecular biology techniques. Leading the way has been polymerase chain reaction (PCR) detection methods. The favorite target for PCR

detection of *Pectinatus* isolates has been the 16S rRNA gene. Detection with species-specific primers (Satokari *et al.*, 1997), PCR followed by colorimetric microplate hybridization (verses traditional detection of PCR products on agarose gels) (Satokari *et al.*, 1998) and growth of the isolates with pre-enriched media before PCR (Juvonen *et al.*, 1999) have all been successful using the 16S rRNA gene region. The detection of the *Pectinatus* 16S rRNA gene has also been extended to fluorescent *in situ* hybridization (commonly referred to as FISH) techniques to limit detection to viable organisms only (Yasuhara *et al.*, 2001). Despite the positive results obtained using this gene region, there is concern that the 16S rRNA gene might be too conserved across bacterial species to allow for specific detection. Satokari *et al.* (1997) did note that the “*Pectinatus*-specific” 16S primers also detected the closely related genus *Zymophilus*. A PCR detection strategy based on the less conserved 16S-23S interspacer region from *Pectinatus* has been established to reduce this cross-reactivity problem (Motoyama and Ogata, 2000b).

With the DNA-based methods to date, a detection level of approximately 10 cells per ml of beer is achieved (Yasuhara *et al.*, 2001). By comparison, the desired sensitivity in the brewery is to detect from 1-10 cells per container of beer (341 ml). This approximately 3000-fold discrepancy between the desired and actual detection limits clearly show that for DNA-based detection to be ideal, much greater sensitivity is needed.

1.3 Flagella

Flagella provide motility to a diverse range of bacteria, propelling them towards environments that promote their survival (Doetsch and Sjoblad, 1980). The basic

bacterial flagellum has three parts: (i) the basal body; which acts as the motor in the cell membrane, (ii) the hook; which is the adaptor link from the basal body to the filament and (iii) the flagella filament; which is a corkscrew-like propeller composed of thousands of flagellin protein monomers (Doetsch and Soblad, 1980; Wilson and Beveridge, 1993). By using a proton or sodium gradient, the basal body rotates the filament to create the propulsion needed to move through a medium (Atsumi *et al.*, 1992). Despite the diversity of organisms that possess flagella, the bacterial flagellum is conserved at both the genetic and structural level throughout the bacterial superkingdom. The bacterial flagellum must be distinguished from the archaea flagellum in that although the function appears to be the same, no homology of flagellum components has been found between these two domains of life (Thomas *et al.*, 2001).

1.4 Flagellin

1.4.1 Basic characteristics

The number of flagellin genes present in an organism varies. Some bacteria have only one flagellin gene, such as *Borrelia burgdorferi* (Wallich *et al.*, 1990), *Escherichia coli* (Ku wajima *et al.*, 1986), *Riftia pachyptila* (Millikan *et al.*, 1999) and *Roseburia cecicola* (Martin and Savage, 1988). Other bacteria have been shown to possess multiple flagellin genes, including *Bacillus thuringiensis* subsp. *alesti* (Lovgren *et al.*, 1993), *Campylobacter jejuni* (Nuijten *et al.*, 1990) and *Salmonella typhimurium* (Homma *et al.*, 1987). An organism will have multiple flagellin genes for different reasons. For some organisms, like *Campylobacter*, both its FlaA and FlaB flagellin subunits are expressed simultaneously and are required together for full motility (Guerry *et al.*, 1990; Winstanely and Morgan, 1997). Other organisms, like *S. typhimurium*,

express only one flagellin subunit at a time and switch between subunits to create phenotypic variation (Homma *et al.*, 1987). This phenomenon is believed to have evolved as a mechanism for pathogenic organisms to avoid the immune response of their hosts.

Flagellin proteins share a characteristic amino acid composition. Amino acids that are regularly found in secondary structure elements, such as valine, leucine and isoleucine, are common in flagellin proteins and alanine, known to be a strong α -helix-forming amino acid, is very common. Conversely, amino acids that disrupt secondary structure are rare. These rare amino acids include proline (which introduces sharp bends), histidine (which can have a variable charge close to neutral pH) and tryptophan (which has a bulky side group) (Wilson and Beveridge, 1993). Cysteine is also very rare, found only as a single residue in the flagellin from *R. cecicola* and as three residues in flagellin from *Pseudomonas putida* (Winstanley and Morgan, 1997). In either case, there is no evidence of disulphide bridge formation involving these residues. This conserved composition of flagellin proteins has been demonstrated repeatedly (Lovgren *et al.*, 1993; Hales *et al.*, 1998).

The structure of the flagella filament is based on the properties of the flagellin protein. Flagellin proteins are comprised of three regions; an N-terminal domain, a central variable region and a C-terminal domain (Figure 1). There are no boundaries separating the three regions, but the N- and C-terminal domains are usually represented as the first 100 amino acids (aa) and the last 50 aa of the protein, respectively (Winstanley and Morgan, 1997). The central variable region makes up the remainder of the protein and accounts for a size variation from 276 aa to 575 aa between flagellin proteins of different species (Wilson and Beveridge, 1993). Flagellin subunits

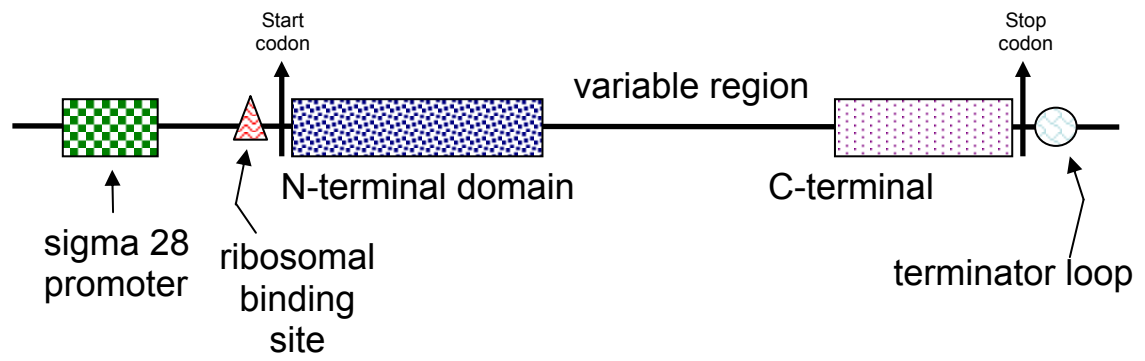


Figure 1 – Generic structure of a flagellin gene.

polymerize in such a way that the N- and C-terminal domains interact to form coiled coils, either within the same flagellin monomer or between adjacent monomers (Wilson and Beveridge, 1993). This domain interaction creates an overall hairpin-like conformation to each flagellin subunit that leaves the central region of each protein exposed to the environment (Homma *et al.*, 1987). As such, the central region of the flagellin is the only part of the protein exposed for antibody binding. This structure explains why the N- and C-terminal regions have remained conserved throughout evolution where as the central region has diverged to create antigenic variability (Winstanley and Morgan, 1997). Deletion studies show that mutations in either the N- or C-terminal domain cause structure deformities to the flagella filament or do not allow for filament assembly at all (Homma *et al.*, 1987; Mimori-Kiyosue *et al.*, 1996). Because the central region of the protein has no structural function, it can be removed without disrupting the assembled flagella. Kuwajima (1988) deleted the central 187 aa from the *E. coli* K-12 flagellin protein (total size of 497 aa) without any loss of function to the flagella. The final filament structure is created by the stacking of folded flagellin monomers into a helix formation that ultimately gives the length of the filament a corkscrew macrostructure (Wilson and Beveridge, 1993).

1.4.2 Regulatory elements

A number of elements must be present in the DNA sequence surrounding a flagellin gene to control its proper transcription and translation. From upstream to downstream, these elements include a sigma 28 promoter, a ribosomal binding site and a terminating loop. These features are shown in Figure 1.

For the initial transcription of any gene, a sigma subunit of bacterial RNA polymerase is needed to recognize an upstream promoter sequence (Losick and Pero, 1976). These sigma factors bind DNA sequences at -35 and -10 base pairs (bp) from the transcriptional start site and, depending on their affinity and abundance, allow for the regulated expression of their genes. The first evidence that flagellin genes had a unique sigma factor to control their transcription was presented by Gilman *et al.* (1981). They showed that *Bacillus subtilis* has a minor sigma factor 28 that recognizes a novel promoter site (CTAAA-16-CCGATAT). This new promoter was shown to regulate the *B. subtilis* flagellin (*hag*) gene as well as flagellar and chemotaxis genes in *E. coli* and *S. typhimurim* (Helmann and Chamberlin, 1987; Mirel and Chamberlin, 1989; Ide *et al.*, 1999). With minor sequence variations, the sigma 28 promoter is present before the flagellin genes in the genera *Campylobacter* (Nuijten *et al.*, 1990), *Clostridium* (Tasteyre *et al.*, 2000), *Rhizobium* (Scharf *et al.*, 2001), *Rhodobacter* (Shah *et al.*, 2000), *Riftia* (Millikan *et al.*, 1999), *Treponema* (Limberger *et al.*, 1992) and *Vibro* (McCarter, 1995) to name a few.

Transcription can be terminated by one of two methods; rho-dependent or rho-independent. Both systems can involve the presence of an RNA stem-loop structure at the end of the transcript. In rho-dependent termination, a rho factor binds the messenger RNA (mRNA) at a rho utilization site and moves up the transcript after the RNA polymerase. When the RNA polymerase stalls at the termination site (for example, at a stem-loop structure), the rho factor catches up and causes the polymerase to release from the DNA template (Snyder and Champness, 1997). In rho-independent termination, also known as intrinsic termination, the mRNA transcript ends with an RNA G+C-rich stem-loop structure followed by a tract of uridine residues. The RNA polymerase is

stalled at the stem-loop structure and the poly-uridine track destabilizes the polymerase to the point where it falls off the DNA template (Yarnell and Roberts, 1999). It has not been reported in the literature whether flagellin genes have a preference for rho-dependent or rho-independent termination. However, a stem-loop structure can be utilized for either method and one example of such a structure is found at the end of the flagellin gene in *B. burgdorferi* (Wallich *et al.*, 1990).

Once a gene has been transcribed into mRNA, a ribosome translates it into protein. Bacterial ribosomes recognize ribosome binding sites or Shine-Dalgarno sequences preceding the start codon by approximately 5-10 bp. These sequences allow the 16S rRNA subunit of the ribosome to bind the transcript and position itself for translation. The traditional Shine-Dalgarno sequence is AGGAGGU (Snyder and Champness, 1997). Flagellin genes have been found to have abbreviated ribosomal binding site that consist of GAGG before *Borrelia* species (Wallich *et al.*, 1990), GGAG before *Pseudomonas* and *Vibro* species (Spangenberg *et al.*, 1996; McGee *et al.*, 1996) or AGGA before *Campylobacter* and *Escherichia* species (Guerry *et al.*, 1990; Kuwajima *et al.*, 1986).

1.4.3 Post-translational modifications

When analyzing a protein, molecular weight can be determined by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) where migration of the protein is compared to the migration of known molecular weight protein standards. This method produces an approximation of the size of the protein and can be correlated with a theoretical protein size determined from a gene sequence. Many flagellin proteins appear significantly larger on SDS-PAGE gel than their corresponding gene sequences

would predict (Wieland *et al.*, 1985; Brimer and Montie, 1998). This size discrepancy is due to post-translational modifications of the flagellin protein. Four types of modifications have been described in the literature pertaining to flagellin. These are glycosylation, phosphorylation, sulfation and the covering of flagella with a sheath.

1.4.3.1 Glycosylation

Glycosylation is the process of adding carbohydrate residues to a protein backbone to create a glycoprotein. There has been extensive work done in the area of eukaryotic glycosylation and comparably little attention paid to the relatively rare cases of prokaryotic glycosylation. Carbohydrate can be linked to proteins through either N- or O-linkages. The N-glycosyl linkage is characterized by the addition of the sugars to the asparagine (Asn) residue in the consensus sequence Asn-Xaa-Ser/Thr (where Xaa can be any amino acid, Ser is a serine and Thr is a threonine) (Lechner and Wieland, 1989). By contrast, O-glycosyl linkages do not have a consensus site, but instead link sugars to Ser or Thr residues that are surface-exposed. Work is continuing to determine what surface features and secondary structure elements are needed to mediate this type of glycosylation (Hansen *et al.*, 1998). Bacterial flagella have been found to contain both types of linkages, with *Halobacter* species being an example of N-glycosylation (Wieland *et al.*, 1985) and *Campylobacter jejuni* being an example of O-glycosylation (Thibault *et al.*, 2001).

Examples in the literature suggest that not all glycosylated flagellin proteins can be detected by current methods. One such case was a study of flagellin proteins from three stains of *P. aeruginosa*. All three strains were closely related and exhibited a comparable size discrepancy between predicted and observed flagellin proteins.

However, glycosylation of only two of the flagellin could be experimentally proven (Brimer and Montie, 1998). This implies that there are unique carbohydrates in bacterial glycoproteins that are undetectable by current methodologies.

1.4.3.2 Phosphorylation

Phosphorylation is the process of adding a phosphate group to a protein. This type of modification is a common regulation method for intercellular proteins. However, it is not a common modification for flagellin proteins. An example of the phosphorylation of a flagellin protein can be found in the a- and b-type flagellin proteins from *P. aeruginosa* (Kelly-Wintenberg *et al.*, 1993).

1.4.3.3 Sulfation

Sulfation is the process of permanently adding a sulfate group to a tyrosine residue. Although rare and generally poorly understood, tyrosine-sulfated proteins have been described in species throughout the plant and animal kingdoms (Moore, 2003). In relation to prokaryotes, *Halobacteria* flagellin proteins are sulfated glycoproteins. This example of sulfation is different from typical sulfation in that the sulfate group is attached to glucuronic acid residues on the glycosylated flagellin instead of tyrosine residues (Wieland *et al.*, 1985). Whether this linkage represents a new type of sulfation or a unique situation to *Halobacteria* remains to be determined.

1.4.3.4 Flagella sheaths

Although not a true post-translational modification, some completed flagella are coated in a sheath structure. Sheathed flagella are common among spirochetes, but

relatively rare throughout other bacterial groups. Exceptions to this include *Helicobacter pylori* and *Vibrio parahaemolyticus* (Geis *et al.*, 1993; McCarter, 1995). Most sheaths appear to be an LPS-like layer similar to the outer membrane (Wilson and Beveridge, 1993). A notable exception to this is for the spirochete *Borrelia*, which appears to have a distinct protein (FlaA) that acts as a sheath (Panelius *et al.*, 2001). The purpose or function of the sheath has not yet been determined.

1.5 Research objectives

Detection of *Pectinatus* isolates in the brewery setting is limited by two major factors. The first limiting factor is that classical detection techniques are generally unsuitable for *Pectinatus* isolates because they are either inappropriate for the growing requirements of *Pectinatus*, cannot offer detection specificity or require specialized reagents or equipment not typically found in a brewery. The second limiting factor is that a lack of knowledge at a genetic level limits DNA-based molecular methods to universally conserved regions of the genome (16S rRNA or 16S-23S interspacer regions). As such, these methods will continue to pose a risk of cross-reactivity with closely related organisms. The flagellin gene of *Pectinatus* offers a new avenue for detection. Flagellin genes contain both conserved and variable regions that can be exploited for a range of detection specificities, conceivably from any flagellated organism down through genus-, species-, subspecies- and isolate-specific levels.

The working hypothesis for this study was that by understanding the *Pectinatus* flagellin gene and protein, new detection protocols based on the flagellin gene or on the flagellin protein could be designed and implemented for brewery quality control of this organism. With that in mind, this project has two main goals. The first is to identify and

characterize the flagellin gene (and the expressed protein there from) of both species of *Pectinatus*. This data gives basic scientific information on these unique organisms and allows for a number of taxonomic-related, phylogenetic analyses to be conducted. The second goal is to investigate whether the flagellin gene would be a suitable marker for rapid protocols (PCR-based) for the detection and identification of *Pectinatus* in the brewery setting. To be suitable as a target in such assays, the *Pectinatus* flagellin gene has to allow for the specific detection of only *Pectinatus* isolates to the exclusion of all other beer spoilage bacteria.

2.0 MATERIALS AND METHODS

2.1 Propagation of bacteria

Pectinatus isolates used throughout this study are listed in Table 1. Flagellin sequences obtained from GenBank for use in this study are listed in Table 2. Beer-spoilage bacteria used to test *Pectinatus* genus- and species-specific PCR primers are listed in Table 3. *Pectinatus* isolates were propagated in MRS (de Man, Rogosa, Sharpe) broth (De Man *et al.*, 1960). To maintain as anaerobic an environment as possible, 5-7 ml tubes were filled with MRS broth, leaving an approximately 1 cm headspace. Tubes were capped tightly immediately after autoclaving and cooled. Samples were inoculated into the bottom third of the tube and recapped as quickly as possible. Incubations were done at 30⁰C. Cultures were maintained by weekly subculture.

Lactobacillus, *Megasphaera*, *Pediococcus* and *Zymomonas* isolates were propagated the same way as *Pectinatus* isolates. *Acetobacter* and *Gluconobacter* isolates were propagated aerobically in Mannitol broth. *Citrobacter*, *Enterobacter*, *Klebsiella*, *Obesumbacterium*, *Proteus* and *Serratia* isolates were propagated aerobically in Nutrient broth. All cultures were incubated at 30⁰C.

To maintain culture purity, isolates were streaked onto agar plates of their corresponding media and incubated aerobically or anaerobically (in Gas-Pak jars) as appropriate. Three representative colonies were picked, combined and grown for use.

Table 1 – *Pectinatus* isolates.

Lab Designation	Species	Source
Pc ATCC	<i>Pectinatus cerevisiiphilus</i>	ATCC 29359 ^{T,a}
Pc DSM	<i>Pectinatus cerevisiiphilus</i>	DSM 20466 ^b
Pc VTT	<i>Pectinatus cerevisiiphilus</i>	VTT-E-81132 ^c
Pf ATCC	<i>Pectinatus frisingensis</i>	ATCC 33332 ^T
Pf DSM	<i>Pectinatus frisingensis</i>	DSM 20465
Pf VTT	<i>Pectinatus frisingensis</i>	VTT-E-80121

^T Type strain.

^a ATCC, American Type Culture Collection, Manassas, Virginia, USA.

^b DSM, German Collection of Microorganisms and Cell Cultures, Braunschweig, Germany.

^c VTT, Technical Research Centre of Finland, Espoo, Finland.

Table 2 – Flagellin gene and protein sequences obtained from GenBank^a.

Species	Nucleotide Accession Number	Protein Accession Number	Taxonomic Phylum ^b
<i>Agrobacterium tumefaciens (flaA)</i> ^c	AE00789	CAA65300.1	Alphaproteobacteria
<i>Azospirillum brasilense (lafI)</i> ^c	U26679	AAA86521.1	Alphaproteobacteria
<i>Aquifex pyrophilus</i> ^c	U17575	AAA88923.1	Aquificae
<i>Bacillus subtilis</i> ^{c,d}	M26947	AAA22437.1	Firmicutes
<i>Bacillus thuringiensis (flaB)</i> ^{c,d}	X67139	CAA47620.1	Firmicutes
<i>Bartonella bacilliformis</i> ^c	L20677	AAA22899.1	Alphaproteobacteria
<i>Bordetella bronchiseptica</i> ^c	L13034	AAA22977.1	Betaproteobacteria
<i>Borrelia burgdorferi</i> ^c	X16833	CAA34735.1	Spirochaetes
<i>Brachyspira hyodysenteriae</i> ^c	X63513	CAA45081.1	Spirochaetes
<i>Burkholderia cepacia</i> ^c	AF011370	AAC38199.1	Betaproteobacteria
<i>Butyrivibrio fibrisolvens</i> ^d	AF026812	AAB82613.1	Firmicutes
<i>Campylobacter jejuni (flaA)</i> ^c	J05635	AAA23024.1	Epsilonproteobacteria
<i>Caulobacter crescentus</i> ^c	J01556	NP_421770.1	Alphaproteobacteria
<i>Clostridium botulinum</i> ^d	AF461538	AAM75948.1	Firmicutes
<i>Clostridium chauvoei</i> ^d	AB058932	BAB87728.1	Firmicutes
<i>Clostridium novyi</i> ^d	AB058936	BAB87734.1	Firmicutes
<i>Clostridium septicum</i> ^d	AB058933	BAB87729.1	Firmicutes
<i>Escherichia coli</i> ^c	M14358	AAA23950.1	Gammaproteobacteria
<i>Helicobacter pylori (flaA)</i> ^c	X60746	CAA43148.1	Epsilonproteobacteria
<i>Legionella pneumophila</i> ^c	X83232	CAA58234.1	Gammaproteobacteria
<i>Listeria monocytogenes</i> ^{c,d}	X65624	CAA46578.1	Firmicutes
<i>Proteus mirabilis (fliC)</i> ^c	L07270	AAA62396.1	Gammaproteobacteria
<i>Pseudomonas aeruginosa</i> ^c	M57501	AAA63458.1	Gammaproteobacteria
<i>Roseburia cecicola</i> ^{c,d}	M20983	AAA26411.1	Firmicutes
<i>Salmonella typhimurium</i> ^c	D13689	BAA02846.1	Gammaproteobacteria
<i>Serratia marcescens</i> ^c	M27219	AAA26556.1	Gammaproteobacteria
<i>Shigella boydii</i> (cryptic gene) ^c	D26165	BAA05151.1	Gammaproteobacteria
<i>Sinorhizobium meliloti (flaA)</i> ^c	M24526	AAA26277.1	Alphaproteobacteria
<i>Treponema phagedenis</i> ^c	M94015	AAC28741.1	Spirochaetes
<i>Vibrio parahaemolyticus (flaA)</i> ^c	AF069392	AAC27801.1	Gammaproteobacteria
<i>Wolinella succinogenes</i> ^c	M82917	AAA61598.1	Epsilonproteobacteria
<i>Yersinia enterocolitica</i> ^c	L33467	AAA93103.1	Gammaproteobacteria

^a Genbank sequences available from Genbank database at <http://www.ncbi.nlm.nih.gov/entrez/>.

^b From NCBI Entrez Taxonomy at

<http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=Taxonomy>.

^c DNA sequences used to design PCR primer Cterm1.

^d DNA sequences used to design PCR primer Cterm2.

Table 3 – Beer spoilage isolates used for genus- and species-specific PCR analysis.

Lab Designation	Species	Source
Ace7	<i>Acetobacter</i> sp.	BSO 7 ^a
Ace8	<i>Acetobacter</i> sp.	BSO 8
Cf	<i>Citrobacter freundii</i>	RUH isolate ^b
Ea	<i>Enterobacter agglomerans</i>	Ingeldew (#127 from wort) ^c
Go	<i>Gluconobacter oxydans</i> subsp. <i>oxydans</i>	ATCC 19357 ^{T,d}
Kp	<i>Klebsiella pneumoniae</i>	RUH isolate
L02	<i>Lactobacillus brevis</i>	CCC B1203 ^e
L14	<i>Lactobacillus delbrueckii</i>	CCC B1240
L20	<i>Lactobacillus paracasei</i>	Ingeldew (I3)
L21	<i>Lactobacillus plantarum</i>	Ingeldew (18A)
Mc	<i>Megasphaera cerevisiae</i>	ATCC 43254 ^T
Op	<i>Obesumbacterium proteus</i> (renamed to <i>Flavobacterium proteus</i>)	ATCC 12841 ^T
P06	<i>Pediococcus clausenii</i>	Molson B71 ^f
P12	<i>Pediococcus damnosus</i>	ATCC 25248
P24	<i>Pediococcus pentosaceus</i>	ATCC 33314 ^T
P25	<i>Pediococcus acidilactici</i>	ATCC 12697
Pm	<i>Proteus mirabilis</i>	RUH isolate
Sm	<i>Serratia marcescens</i>	Dairy isolate ^g
Za	<i>Zymomonas anaerobia</i>	BSO 57
Za [sub a]	<i>Zymomonas anaerobia</i> subsp. <i>anaerobia</i>	NCIMB 8227 ^h

^T Type strain.

^a BSO, Beer Spoilage Organism, Dr. B. Kirsop, Institute for Biotechnology, Cambridge, England; obtained through Dr. W. M. Ingeldew, College of Agriculture, University of Saskatchewan, Saskatoon, Saskatchewan, Canada.

^b RUH, Royal University Hospital, Saskatoon, Saskatchewan, Canada.

^c Dr. W. M. Ingeldew, College of Agriculture, University of Saskatchewan, Saskatoon, Saskatchewan, Canada.

^d ATCC, American Type Culture Collection, Manassas, Virginia, USA.

^e CCC, Coors Brewing Company, Golden, Colorado, USA.

^f Molson, Molson Breweries of Canada Limited, Montreal, Quebec, Canada; now deposited as ATCC BA-344^T.

^g Microbiology Teaching Laboratory, Department of Microbiology and Immunology, University of Saskatchewan, Saskatoon, Saskatchewan, Canada.

^h NCIMB, National Collection of Industrial and Marine Bacteria, Aberdeen, Scotland.

2.2 Isolation of flagellin protein

Flagellin protein used in this study was extracted in late 1992 by Ms. Jillian Hymers and Dr. Barry Ziola. Briefly, the procedure used was as follows. One hundred ml of each *Pectinatus* culture was grown and subjected to centrifugation at 4°C in a Sorvall RC-5B Refrigerated Superspeed Centrifuge (SA 600 rotor) for 10 min at 11,000 g. Pellets were resuspended with 30 ml of sterile phosphate-buffered saline (PBS). The bacterial cells were deflagellated by shearing with a Sorvall Omnimixer at room temperature for 80 sec. Samples were cooled on ice and centrifuged at 4°C in the Sorvall RC-5B for 15 min at 11,000 g. The supernatants were saved and pellets were resuspended in 2.5 ml of sterile PBS and centrifuged again. Supernatants were pooled and the flagella centrifuged at 4°C in a Beckman L8-70M Ultracentrifuge at 100,000 g for 1 hr. Flagella pellets were resuspended in 1.6 ml of sterile PBS and stored at -20°C. Protein determinations were done using the BCA Protein Assay. Absorbance was read with a Bio-Rad Model 2550 EIA Reader at 565 nm and interpreted by Bio-Rad MacReader II software.

2.3 Preparation of flagellin protein for N-terminal sequencing

The full-length flagellin subunits and peptides derived from the flagellin were isolated for N-terminal sequencing. To isolate the full-length proteins, 48 µg and 41 µg of *P. cerevisiophilus* (Pc) ATCC and *P. frisingensis* (Pf) ATCC flagellin, respectively, were used. To the flagellin samples, 94 µl of SDS-PAGE dissociation solution were added and the entire sample boiled for 5 min. Flagellin were separated from minor contaminants by SDS-PAGE (section 2.4).

Alpha-chymotrypsin was used to cleave flagellin subunits to obtain peptides for N-terminal sequencing. For Pc ATCC and Pf ATCC, 90 µg of flagellin protein was digested with 4.5 µg of α-chymotrypsin (in reverse osmosed, deionized water (rdH₂O)) for 15 min or 2.25 µg of α-chymotrypsin (in rdH₂O) for 2.5 min, respectively. The digestion was stopped by the addition of 40 µl of SDS-PAGE dissociation solution and boiling for 5 min. Peptide fragments were separated by SDS-PAGE (section 2.4).

After SDS-PAGE, proteins were transferred to polyvinylidene difluoride (PVDF) membrane (section 2.12). Membranes were visualized with BLOT-FastStain. Desired bands were cut out with a scalpel, destained with rdH₂O and dried.

2.4 Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE)

Samples were prepared as individual sections specify. Proteins were separated by SDS-PAGE using a discontinuous buffer system (Laemmli, 1970), with 1.0-mm thick 12.5% (w/v) acrylamide (29:1 acrylamide-bisacrylamide) separating gels, and 4% (w/v) stacking gels. A 10-well comb was used. Prestained protein molecular weight standards, high range were used in the marker lane. Electrophoresis was carried out at 30 mA constant current (Bio-Rad Laboratories power supply 1000/500) until the tracking dye front reached the bottom of the gel (approximately 4 hr).

2.5 N-terminal amino acid micro-sequencing

N-terminal amino acid sequencing was done at the University of Victoria Protein Chemistry Centre by Ms. Sandy Kielland and Ms. Laura Aitken. The protein sequence analysis was performed using pulsed-liquid phase (473) and gas phase (475) ABI sequencers.

2.6 DNA extraction

Bacterial DNA extractions were done using the Puregene Genomic DNA purification kit. The manufacturer's protocol, "DNA Purification from 0.5 ml Gram-Negative Bacterial Culture" was used. The only deviation from the stated procedure was that 1.5 ml of culture (instead of 0.5 ml) was used as the starting culture volume. Rehydrated DNA was stored at -20°C.

2.7 Polymerase chain reaction (PCR)

2.7.1 Reaction components

The standard PCR mixture used contained 1x PCR buffer, 0.2 mM of each of the four deoxynucleotide triphosphates (dNTPs), 0.2 µM of each primer (section 2.7.2) and 4 units of Platinum *Taq* polymerase. Either 1 µl of genomic DNA (rehydrated as per manufacturer's instructions, section 2.6) or 5 µl (approximately 2.5 ng) of Bubble-linked DNA was added as the template DNA and the final volume of each reaction tube was brought up to 50 µl with sterile rdH₂O. Reactions were run in either a PTC-100 Programmable Thermal Controller or a PCR Sprint Temperature Cycling System. At the end of the amplification program, reactions were cooled to 4°C and visualized on an agarose gel or transferred to -20°C for storage. Each set of samples included a negative control consisting of all the reaction components but with template DNA replaced by either water that had been treated through the DNA extraction procedure (when genomic DNA was used as the template) or water that had been "ligated" with Bubble-linkers (when Bubble-linked DNA was used as the template). The negative control had to contain no amplification products for the PCR results for a set of samples to be considered valid.

2.7.2 Primer design

A variety of different PCR primers were used throughout this study (Table 4). With the exceptions of the primers 8F and 534R, which were taken from Dobson *et al.* (2002), and BubblePrimer (and by extension, SeqBubble), which was taken from Munroe *et al.* (1994), PCR primers were designed as needed. PCR primers designed in this study can be divided into four general groups; primers designed from amino acid sequences, primers designed to specific DNA sequences for Bubble-PCR, a primer set designed to amplify the full-length flagellin gene for cloning and primer sets designed to specifically detect the *Pectinatus* flagellin gene at the genus or species level. Each group of primers required a different set of design considerations.

In this study, the primers FlgnTerminal1, Flgn2F, Flgn3R, Flgn9F, Flgn10F, Cterm1 and Cterm2 were designed from amino acid sequences. All of the Flgn primers are designed from portions of the flagellin aa sequence from Pc ATCC. Primers FlgnTerminal1, Flgn2F, Flgn3R and Flgn10F are the result of converting a short amino acid sequence into its corresponding DNA sequences, allowing for codon degeneracy, and creating a single, degenerate PCR primer that has the ability to bind all the possible DNA sequences. Flgn9F is a specific primer that was developed from the DNA sequence that codes for the aa sequence used to design primers Flgn2F and Flgn10F. Cterm1 is a PCR primer designed to bind within the C-terminus domain of all bacterial flagellin genes. For its design, 27 known flagellin sequences (Table 2) were first aligned according to their protein sequences. From the protein alignment, conserved regions were identified within the C-terminus and the corresponding DNA sequences for these regions were obtained. A degenerate consensus sequence was determined. Primer Cterm1 was designed to anneal to this consensus sequence. Primer Cterm2 was

Table 4 – PCR primers.

Primer Name	Sequence (5'-3')	Designed For	Primer Binding Location ^a
FlgnTerminal1	CTNGTNGTNAARAAYAAAYATGTC	Pc	256-278
Flgn2F	GCNYTNGTNGTNAARAAYAAAYATG	Pc	253-276
Flgn3R	CATNSWNCCRTCACAAARAARYT	Pc	649-672
Flgn4R	CATAACCTGCTGGATTGTCTGC	Pc	373-394
Flgn5F	TGACCGCAATACCATTTCAGGAT	Pc	570-591
Flgn6F	AGTGAACAGATTAGCGCCCAGG	Pc	1141-1162
Flgn7F	CTAAAGCAAACCACTCCATTTCG	Pc	1202-1223
Flgn8F	CTACTCGTGCTAATGCTAATGC	Pc	1307-1328
Flgn9F	GCTTTAGTAGTAAAAACAACATG	Pc	253-276
Flgn10F	GCIYTIGTIGTIAARAAYAAAYATG	Pc	253-276
Flgn11F	TGCTAATATGGCTCAGGAAATG	Pf	1476-1497
Flgn12R	CGTTGTCAAGTTATCCTGTGTA	Pf	1413-1434
Flgn13R	CGTTAATCTGGTCAACTGCCTG	Pf	601-622
Flgn14R ^b	TTTCTGACCAGATGATACCTTC	Pf	339-360
Cterm1	YTGRTTIGCYTGIGCMAGCA	Both	1544-1564
Cterm2	GTGGTTGTTGATTAGCTTGAGC	Both	1549-1570
PectCloneF	<u>CGCGGATCCATGGCTTTAGTAGTAAAAAY</u>	Both	250-270
PectCloneR	<u>GGCCGTCGACMMATTATGCAACAACTGC</u>	Both	1581-1599
Pect550F	ATGCWGCCAATGATACCAATAC	Both	542-563
Pect550R	CCAGTKATGCTGATAGTAAAACC	Both	1069-1091
Pf253F	CAGGTACAGCATTTTCTTCCAG	Pf	725-742
Pf253R	AACCCGATAAAGCTTCCACTACC	Pf	955-977
Pc185F	AGCGACTGATTTTGGTTCTGC	Pc	726-746
Pc185R	ACCCTGTAGCACTTGTAACACTG	Pc	888-910
8F	AGAGTTTGATCCTGGCTCAG	All 16S	8-27
534R	ATTACCGCGGCTGCTGG	All 16S	518-534
BubblePrimer	GCGGCCGCGAGTTCGTCAACATAGCATTTCT	-	-
SeqBubble	CCGCAGTTCGTCAACATAGC	-	-

R=A+G; Y=C+T; M=A+C; S=G+C; W=A+T; N=A+T+C+G; I=inosine

All primers were synthesized by Sigma-Genosys and purified by desalting. Underlined sections of PectCloneF and PectCloneR indicate restriction sites added.

^a Numbers refer to sequence numbering in Figure 8. Primers designed to Pc or Both are referenced to flagellin sequence Pc ATCC; primers designed to Pf are referenced to flagellin sequence Pf DSM. Primers 8F and 534R are referenced to the *E. coli* 16S rRNA gene numbering system.

^b Flgn14R was used in a PCR reaction mixture containing 2.5 mM MgCl₂. All other primers were used in a PCR reaction mixture containing 1.5 mM MgCl₂.

designed the same way as primer Cterm1; however, its specificity is for flagellin genes from members of the phylum *Firmicutes*. Primer Cterm2 was designed using the nine sequences from the phylum *Firmicutes* (Table 2).

Flgn4 to Flgn8 and Flgn11 to Flgn14 are specific PCR primers designed for Bubble-PCR. Many properties were assessed for these primers. The first consideration was the properties of the primer region. Ideally, primer sequences were from 20 to 23 bp in length, as close to a 50% G+C content as possible, had a C or G as the terminal 3' base and had at least one A or T in the 3' triplet. Primer sequences meeting these criteria were checked for secondary structure elements and the ability to form primer dimers with the program DNA calculator (Sigma website; <http://www.sigma-genosys.com/>). DNA calculator also determined a melting temperature for the primers, which needed to be above 50°C. Acceptable primer characteristics were weak to no secondary structure and no ability to form primer dimers. Sequences having acceptable PCR primer characteristics were checked for possible alternative binding sites by comparison to sequences in GenBank. The Basic Local Alignment Search Tool (BLAST) option Nucleotide: Search for short, nearly exact matches was used, limited to sequences of bacterial origin (Altschul *et al.*, 1990). Regions of DNA having no significant sequence similarity (less than 16 bp match) to any known bacterial sequence were designated as acceptable. All specific PCR primers used in this study met these conditions.

The PCR primer set, PectCloneF and PectCloneR, were designed to amplify the full-length flagellin gene from *Pectinatus* isolates and allow for the product to be inserted into an expression vector. To do this, the PCR product generated had to have three characteristics. First, the resulting PCR product had to include the entire open

reading frame of the flagellin gene with as little sequence before the start codon and after the stop codon as possible. Second, the restriction sites BamHI and SalI had to be inserted at the beginning and end of the sequence, respectively, along with enough of a nucleotide overhang on each side of the restriction site to allow for cleavage. Third, the BamHI restriction site had to be in-frame with the vector tag so that when the PCR product was cloned and expressed, the proper fusion protein was generated. Primers PectCloneF and PectCloneR were designed to create an amplification product that met all three criteria. These primers were also used to sequence the flagellin genes of Pc DSM, Pc VTT, Pf ATCC and Pf VTT.

Three sets of PCR primers were designed to detect three specific groups; Pect550 was designed to detect all *Pectinatus* isolates, Pf253 was designed to detect only Pf isolates and Pc185 was designed to detect only Pc isolates. At the time of primer design, DNA sequences of the flagellin genes from only Pc ATCC and Pf DSM were available. As such, primer design was limited to these two isolates. Pc ATCC and Pf DSM flagellin DNA sequences were aligned with each other (using Needle, contained within the EMBOSS software package) and the N- and C-terminal domains were marked. The alignments were manually inspected for sections of sequence within or adjacent to the variable region that shared either high identity (for genus-specific) or high disparity (for species-specific). Regions with a length greater than 20 bp were tested for suitable PCR primer characteristics as outlined above for specific PCR primers. Sets of forward-reverse primers were chosen to detect each target group. No PCR primer set was permitted to generate a PCR product within 50 bp of another set. Final primer sets were checked with Primer3 (Rozen and Skaletsky, 2000) to assess compatibility, primer properties and product size.

2.7.3 PCR programs

Four different PCR programs were used throughout this study. A generic program was used for sets of degenerate PCR primers and Bubble-PCR. For full-length flagellin gene amplification and genus- and species-specific PCR detection, the annealing temperature and number of cycles in the generic PCR program were altered to optimize conditions. Finally, amplification of the 16S rRNA region was carried out with a published, pre-determined program. The generic PCR program used was 95⁰C for 5 min (initial denaturation), 40 amplification cycles consisting of 94⁰C for 45 sec (denature), 40⁰C (annealing for degenerate primers) or 60⁰C (annealing of specific primers) for 45 sec and 72⁰C for 50 sec (extension), followed by a final extension step at 72⁰C for 5 min. The second PCR program used was to amplify the full-length flagellin genes with the PCR primers, PectCloneF and PectCloneR. The generic PCR program was altered so that a total of 30 amplification cycles were run, with the first 5 cycles carried out at a 50⁰C annealing temperature and the remaining 25 cycles at a 60⁰C annealing temperature. The rest of the program remained unchanged. The third PCR program used the PCR primer sets, Pect550, Pf253 and Pc185, to specifically detect all *Pectinatus* species, only Pf isolates and only Pc isolates, respectively. Alternations to the generic PCR program included reactions being carried out with 30 amplification cycles. As well, an annealing temperature of 59⁰C was used for the Pect550 primer set and 61⁰C used for the Pf253 and Pc185 primer sets. To amplify the 16S rRNA gene, both the reaction mixture composition and the PCR program were taken from Dobson *et al.* (2002). The standard PCR reaction mixture was altered to contain 2.5 units of Platinum *Taq* polymerase, 1 µl of template DNA and the final volume of each tube was brought up to 100 µl. The 16S PCR program consisted of 5 min at 95⁰C, 35 cycles

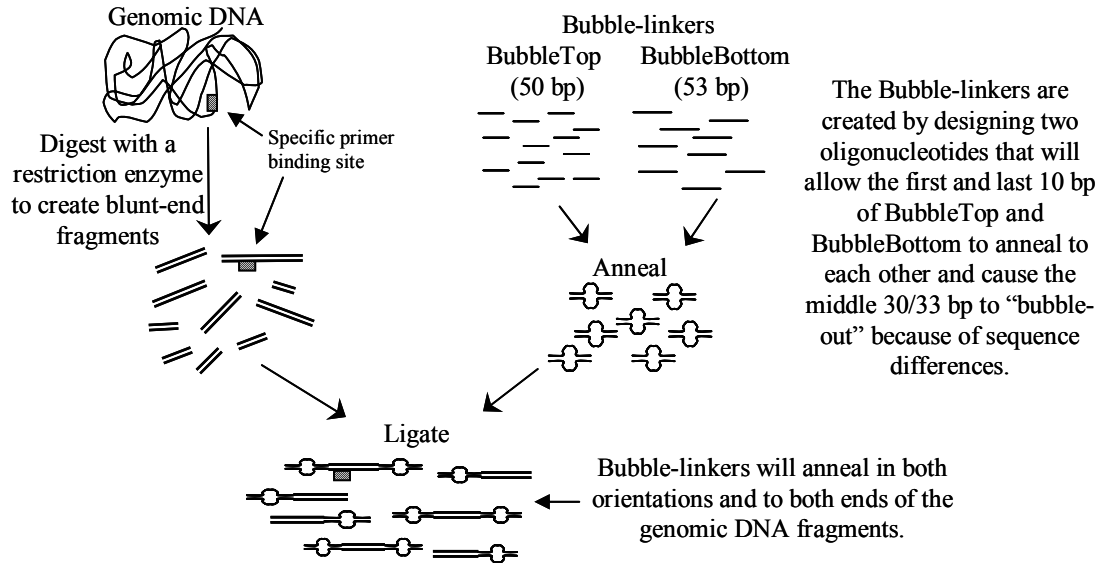
containing 1.25 min at 94⁰C, 1 min at 60⁰C and 1 min at 72⁰C, 10 min at 72⁰C and a final hold at 5⁰C.

2.7.4 Bubble-PCR

Bubble-PCR is a variation of the standard PCR technique. Originally introduced by Riley *et al.* (1990), the technique has also been called vectorette PCR and genome walking. Bubble-PCR is used to amplify a region of DNA with only one specific primer. This can be done because the target DNA is digested with a restriction enzyme and a linker molecule of DNA is ligated onto the ends. The second PCR primer, BubblePrimer, is designed to anneal to the Bubble-linker construct. For this study, three pools of Bubble-linked genomic DNA were prepared; one digested at HaeIII sites, one at RsaI sites and one at EcoRV sites. This created three different locations for BubblePrimer to match up with any specific primer. The procedure is diagramed in Figure 2.

Bubble-linked genomic DNA was prepared as described by Munroe *et al.* (1994). The Bubble-linker was constructed from the oligonucleotides BubbleTop, GAAGGAGAGGACGCTGTCTGTCGAAGGTAAACGGACGAGAGAAGGGAGAG, and BubbleBottom, CTCTCCCTTCTGCGGCCGACGTTTCGTCAACATAGCATTCTGTCCTCTCCTTC, suspended in Tris(hydroxymethyl)aminomethane (Tris)-ethylenediaminetetraacetic acid (EDTA) buffer, pH 7.6 (TE buffer), to a concentration of 100 mM. To prepare the linkers, 9 µl of both BubbleTop and BubbleBottom were added to 282 µl of 2x SSC (saline-sodium citrate)/Tris to make the final concentration of the linkers 3 pmol/µl. The cocktail was boiled for 10 min and allowed to cool to room temperature. The genomic DNA was prepared for linkage by digesting 0.25 µg of DNA

Preparing template DNA:



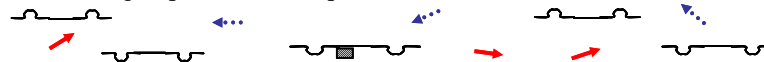
PCR Primers:

Specific primer (→) – A 20-22 bp oligonucleotide to your DNA sequence of interest. This primer is designed so that its 3' end is directed away from the known DNA region.

BubblePrimer (←··) – A 32 bp oligonucleotide that is exactly the same sequence as the middle section of BubbleBottom. Because BubbleTop is not complementary to this region, BubblePrimer will not have a binding site until BubbleBottom is complemented.

In PCR:

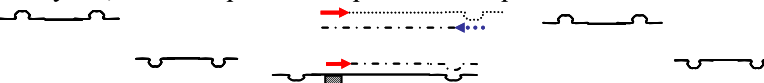
1) The first denaturing step in the PCR separates the DNA strands:



2) The specific primer finds its complementary DNA region and primes DNA replication. A sub-population of the genomic DNA fragments containing the specific primer binding site will have the BubbleBottom DNA strand in the correct orientation to allow it to be complemented:



3) The next PCR cycle denatures the DNA strands again. Now that BubblePrimer has a complementary site, both PCR primers can prime DNA replication:



4) Continued PCR cycling causes only the exponential amplification of the DNA region between the specific primer and BubblePrimer:

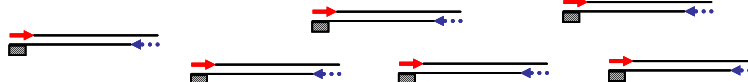


Figure 2 – Diagram of Bubble-PCR methodology.

with 20 units of RsaI, HaeIII or EcoRV, in appropriate buffers, at 37⁰C overnight. The enzymes were heat-inactivated at 65⁰C (for RsaI) or 80⁰C (for HaeIII and EcoRV) for 20 min. Bubble-linked DNA was produced by ligating digested genomic DNA with 12 pmol of annealed Bubble-linkers (4 µl), 2 units of T4 DNA ligase (2 µl) and 20 µl of 5x ligase reaction buffer. The volume of the reaction was made up to 100 µl with TE buffer and incubated overnight at room temperature. For use in PCR, the Bubble-linked DNA was diluted 1:100 with ¼ TE buffer and stored at -20⁰C.

2.8 PCR product visualization and purification

PCR products were visualized on 0.6% agarose gels made with 1x TBE buffer (Tris-boric acid-EDTA) containing 0.5 µg/ml ethidium bromide. For each sample, 15 µl from the PCR reaction was combined with 2.5 µl of tracking dye and loaded onto a gel. A 100 bp ladder was run on each gel. Gels were run in 1x TBE buffer containing 0.5 µg/ml ethidium bromide at 108 volts for approximately 50 min. Bands were then visualized using an ultraviolet (UV) light source.

PCR products chosen for sequencing were purified in one of two ways. If the desired band was the only product in the PCR sample, 85 µl of PCR sample was purified and concentrated with the Wizard PCR Preps DNA Purification System (VWR Canlab, Mississauga, ON) using the procedure for Direct Purification of DNA from PCR Amplification without a vacuum manifold. The only deviation from the stated protocol was that 30 µl of TE buffer was used for the final elution of DNA (instead of 50 µl).

If the desired band was not the only product in the PCR reaction, a cheese-cloth extraction procedure was used. Eighty-five µl of PCR sample was mixed with 15 µl of tracking dye and loaded into 3-4 lanes on a 0.6% agarose gel (as described above). A

0.5 ml thick-walled eppendorf tube was prepared for each sample. A hole was punched in the bottom of the tube and a 2.5 cm² piece of sterile cheesecloth was tightly compacted into the bottom. The 0.5 ml tube was set inside a 1.5 ml eppendorf tube. Using a UV box, the desired band was cut from the gel using a scalpel, placed into the prepared 0.5 ml tube and frozen at -70⁰C for 30 min. The samples were left to stand at room temperature for 5 min and centrifuged on an Eppendorf 5415C table-top centrifuge (all centrifugations were done at top speed) for 10 min. The resulting buffer/DNA mix was collected in the 1.5 ml tube.

The DNA-containing solution was then extracted with a phenol-chloroform procedure to remove any contaminating protein. Briefly, an equivalent volume of phenol-chloroform-isoamyl alcohol (25:24:1) was added, shaken to emulsify, and centrifuged for 3 min. The DNA-containing aqueous layer (top layer) was transferred to a new tube. An equal volume of chloroform-isoamyl alcohol (24:1) was added and the extraction repeated.

DNA was precipitated and concentrated by the addition of 10 µl of 5 M NaCl and 2 volumes of cold 95% ethanol. Samples were inverted gently to mix and incubated at -20⁰C for 30 min. The DNA solution was then centrifuged for 10 min. The ethanol was removed, the pellet rinsed with 70% ethanol and centrifuged for 3 min. The ethanol was again removed and the pellet allowed to air dry. The DNA pellet was then resuspended in 25-30 µl of TE buffer.

Before sequencing, the purified PCR products were quantified. This was done by running 2-5 µl of the purified product on a 0.6% agarose gel (as described above) beside 4 µl of a low DNA mass ladder. The quantity of DNA of the purified PCR product was estimated in comparison to band intensities of the mass ladder.

2.9 DNA sequencing

DNA sequencing was performed at the Plant Biotechnology Institute, National Research Council, Saskatoon, Saskatchewan. Direct sequencing of the PCR products was performed by the fluorescence-based dideoxy termination method with the thermocycler model: GeneAmp PCR system 9700. The sequencing reaction was set up with the Big Dye Terminator kit, using 3.2 pM of primer and the appropriate amount of DNA, in ng, for the size of PCR product to be sequenced. The reaction was run for 25 cycles with 96⁰C for 10 sec, 50⁰C for 5 sec and 60⁰C for 4 min. The samples were then sequenced with the ABI 373 DNA sequencer STRETCH model. For sequencing primers, the original PCR primers for each particular PCR product were used. With the Bubble-PCR generated products, the primer SeqBubble was used in place of BubblePrimer.

2.10 Sequence analysis and assembly

All sequencing results were checked manually for base-calling errors and queried for similarities with BLAST (Altschul *et al.*, 1990). PCR products generating high quality sequence reads were only sequenced in one direction, whereas lower quality sequence reads or short PCR products that required very accurate base-calling at both ends were sequenced in both directions. For the assembly of the complete flagellin sequences of Pc ATCC and Pf DSM, at least 50 base pairs (bp) of high quality sequence overlap was used to join PCR products. From the EMBOSS software package, the programs, REVSEQ, used to reverse-complement sequences, MERGER, used to merge two overlapping sequences together, and SHOWSEQ, used to visualize translations and restriction sites, were utilized. Flagellin sequences and 16S rRNA sequences were

multiply aligned and sequence differences were manually checked from sequencing results.

2.11 Cloning of the flagellin genes

2.11.1 Glutathione S-transferase (GST) gene fusion system

The expression system used in this study was the pGEX plasmid vector from the GST Gene Fusion System (Amersham Biosciences, Baie d'Urfe, QB). The GST gene fusion system allows for the creation and expression of a fusion protein with *Schistosoma japonicum* GST and a gene of interest. The pGEX plasmids are designed for inducible, high-level intracellular expression of fusion proteins. Figure 3 is a map of the vector used in this study, pGEX-6P-1, showing the reading frames and main features.

2.11.2 Cloning procedure

The full-length flagellin genes from Pc ATCC and Pf DSM were amplified with the PCR primers, PectCloneF and PectCloneR, as specified in section 2.7.1.2. PCR products were confirmed and quantified on 0.6% agarose gels (section 2.8).

To prepare the flagellin genes for insertion into the vector, PCR products were purified using the phenol-chloroform, ethanol precipitation procedure (section 2.8). Ten μg of PCR product was digested with 30 units of both BamHI and Sall, using the Sall buffers, at 37°C for 2.5 hr. The mixture was purified with the phenol-chloroform, ethanol precipitation procedure (section 2.8) and resuspended in 20 μl of sterile rdH₂O.

To prepare pGEX-6P-1, 1 μg of vector DNA was brought to a volume of 15 μl with rdH₂O and digested with 30 units of both BamHI and Sall, using Sall buffers, at

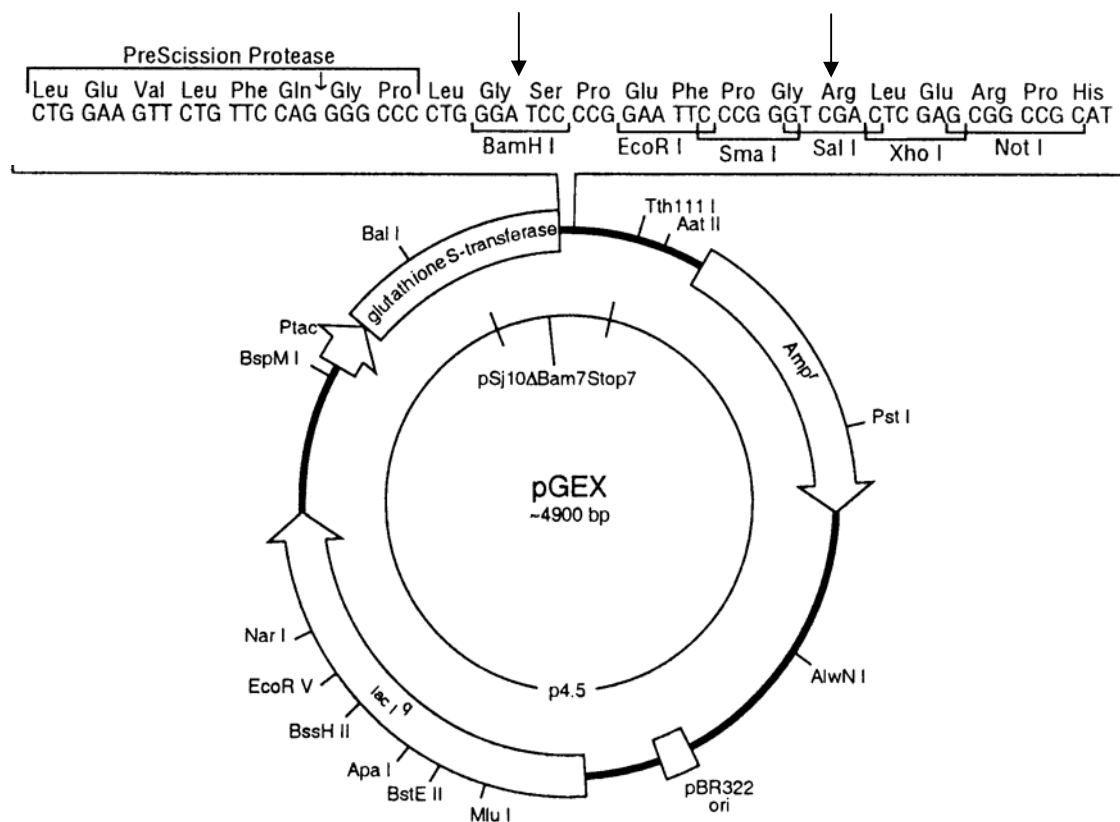


Figure 3 - Map of the glutathione S-transferase fusion vector, pGEX-6P-1, showing the reading frame and main features. Reproduced from page 12 of the Amersham Biosciences GST Gene Fusion System Handbook (2002). Arrows indicate cloning sites used.

37⁰C for 1.5 hr. Fifty units of alkaline phosphatase and 5 µl of 10x alkaline phosphatase buffer were added and the incubation continued for another hour. The digested vector was purified using the phenol-chloroform, ethanol precipitation procedure (section 2.8) and resuspended in 20 µl of rdH₂O.

To assemble the vector, 5 µg of insert DNA (10 µl) and 0.25 µg of pGEX-6P-1 (5 µl) were combined with 1 unit of T4 DNA ligase (1 µl) and 4 µl of 5x ligase buffer and incubated at 16⁰C overnight. The assembled vector was purified by phenol-chloroform, ethanol precipitation (section 2.8), resuspended in 10 µl of sterile rdH₂O and stored at -20⁰C.

Transformations of the vectors pGEX-Pc (pGEX-6P-1 containing the flagellin gene from Pc ATCC), pGEX-Pf (pGEX-6P-1 containing the flagellin gene from Pf DSM) and pGEX (pGEX-6P-1 vector alone as a control) were done with DH10β *E. coli* cells. Predispensed aliquots of competent DH10β *E. coli* cells were gratefully obtained from Dr. Wei Xiao. To begin, *E. coli* cells were thawed on ice and cuvettes and the white slide were cooled on ice. For each vector, 3 µl of vector DNA were mixed into an aliquot of cells. The cells were transferred into a cuvette, tapped to ensure the cells were in the bottom and incubated on ice for 45 sec. Cuvettes were put into the white slide and inserted into the chamber of the *E. coli* Pulser Transformation apparatus (Bio-Rad Laboratories, Mississauga, ON). The cells were pulsed at 1.8 volts with an *E. coli* Pulser (Bio-Rad Laboratories). Four hundred µl of SOC medium were added to each cuvette, mixed and the cell suspension transferred to a 1.5 ml eppendorf tube to be incubated at 37⁰C for 45 min. Finally, 300 µl of the cell suspension were spread onto Luria-Bertani medium + ampicillin (LB+amp) plates and incubated at 37⁰C overnight.

Plates were examined for the appearance of white colonies, indicating successfully transformed cells.

Plasmids were recovered from transformed cells using a bacterial plasmid mini-prep protocol. Transformed cells were picked from LB+amp plates, inoculated into 1.5 ml of LB+amp broth and incubated overnight at 37°C. Cells were collected by centrifugation in an Eppendorf 5415C table-top centrifuge (all centrifugations were done at top speed) for 30 sec and resuspended in 350 µl of bacterial plasmid mini-prep solution (8% sucrose, 0.5% Triton-x-100, 50 mM EDTA (pH 8.0) and 10 mM Tris-HCl (pH 8.0)). Twenty µl of lysozyme (10 mg/ml in 10 mM Tris-HCl, pH 8.0) were added and tubes were inverted to mix. Cells were lysed by boiling for 30-40 sec and centrifuged for 10 min. Cell walls and membranes were removed using a sterile toothpick. The plasmid DNA was concentrated using the ethanol precipitation method (section 2.8) and resuspended in 30 µl of sterile rdH₂O.

The correct assembly of the vectors was confirmed by the ability to recover the flagellin insert from the plasmid extracted and by the overall size of the plasmid. To recover the flagellin insert, 10 µl of the purified plasmid DNA was digested with 20 units of both BamHI and Sall, using the Sall buffers, overnight at 37°C. All of the digested plasmid DNA, as well as 4 µl of the undigested plasmid DNA were run on 0.6% agarose gels with 1x TBE buffer. No ethidium bromide was included in either the gel or the buffer during electrophoresis. The gels were stained after electrophoresis in a solution containing 1 mg of ethidium bromide per 500 ml of water for 10 min and visualized on a UV light source.

2.12 Detection of GST-flagellin fusion proteins

2.12.1 Extraction of fusion proteins

The fusion protein extraction procedure was modified from The GST Gene Fusion Handbook (Amersham Biosciences, 2002). Transformed *E. coli* cells were inoculated into 1.5 ml of LB+amp broth and incubated for 4 hr at 35⁰C, with vigorous agitation. Fusion proteins were induced with 1.5 µl of 100 mM isopropylthio-B-D-galactoside (commonly known as IPTG) and the incubation was continued. At either 1.25 hr or 1.5 hr post-induction, cultures were centrifuged in an Eppendorf 5415C table-top centrifuge (all centrifugations were done at top speed) for 30 sec. The medium was removed and the pellets were resuspended in 175 µl of bacterial plasmid mini-prep solution. Twenty µl of lysozyme were added, mixed and the cells were boiled for 40 sec. For analysis, 90 µl of the whole-cell protein lysate were added to 40 µl of SDS-PAGE dissociation solution and boiled for 10 min. As a positive control, 5 µg of extracted flagellin protein (2.5 µg from Pc DSM and 2.5 µg from Pc VTT) were combined with 40 µl of SDS-PAGE dissociation solution, boiled for 10 min and included with each fusion protein gel. Protein samples were loaded and SDS-PAGE was carried out (section 2.4).

2.12.2 Immunoblotting of fusion proteins

Protein transfer from the separating gel to a PVDF membrane (Bio-Rad Laboratories) was done by diffusion. Separating gels were prepared for transfer by incubation in transfer buffer (25 mM Tris, 192 mM glycine and 20% methanol), with gentle rocking, for 1 hr. PVDF membranes were dipped into methanol and then soaked in transfer buffer with pieces of blotting paper for 1 hr. Gels were then sandwiched

between two membranes, four pieces of blotting paper (two on each side) and two pieces of glass. The whole apparatus was wrapped in saran wrap, then aluminum foil, and left undisturbed for 4 days at room temperature with a 5 kg weight on top. Membranes were removed and washed with Tris-buffered saline (TBS; 20 mM Tris-HCl, pH 7.5, containing 500 mM NaCl) by gentle rocking for 5 min.

Excess binding sites on the membranes were blocked with blocking solution (PBS, containing 3% (w/v) bovine serum albumin (BSA) and 0.5 mM merthiolate) at 4°C overnight. Membranes were then washed for three 5 min intervals by gentle rocking in TTBS (TBS containing 0.05% (v/v) Tween 20). The primary antibody was added and the membranes incubated at room temperature for 2 hr with gentle rocking. For the identification of *Pectinatus* flagellin protein, the primary antibody was mouse polyclonal anti-*Pectinatus* flagellin antiserum (prepared by Ms. Jillian Hymers and Dr. Barry Ziola) diluted 1:400 in antibody buffer (PBS, 2% (w/v) BSA, 2% Tween 20 and 0.2 mM merthiolate). For the identification of GST protein, the primary antibody was goat polyclonal anti-GST antiserum (Amersham Biosciences) diluted 1:50,000 in antibody buffer. After washing as before, the membranes were incubated for 1.5 hr with alkaline phosphatase-conjugated indicator antibodies. AffiniPure F(ab')₂ fragment goat anti-mouse IgG (H+L) (Cedarlane Laboratories Ltd., Hornby, ON) diluted 1:4,000 with antibody buffer was used for *Pectinatus* flagellin detection whereas rabbit anti-goat IgG (H+L) (EMD Biosciences, Inc., San Diego, CA) antibodies diluted 1:2,000 with antibody buffer was used for the detection of GST. Membranes were again washed three times with TTBS and twice with TBS. A color development system (BCIP/NBT; Bio-Rad Laboratories) was used with stock solutions of the two chemicals diluted 1:1,000 in 100 mM Tris-HCl, pH 9.5, containing 0.5 mM MgCl₂. When the desired band

intensity was reached, color development was stopped by rinsing the membranes with water.

2.13 Carbohydrate detection

2.13.1 Carbohydrate-specific stains

Three different staining procedures were used; Periodic Acid-Schiff's (PAS) staining, alcian blue stain and thymol-sulfuric acid staining. These stains were tested against some combination of the six flagellin proteins, ovalbumin (Ova), fetuin (Fet), horse radish peroxidase (HRP) and BSA. With all SDS-PAGE gels, 60 µg of protein per lane was used, with the exception of BSA, which was used at 20 µg of protein per lane. The amount of BSA was reduced so that Coomassie Brilliant Blue staining of the SDS-PAGE gels gave equal staining intensity to each protein band. Protein samples were brought to a volume of 80 µl with rdH₂O and 60 µl of SDS-PAGE dissociation solution was added. The samples were boiled for 5 min and loaded on SDS-PAGE gels (section 2.4). All separating gels were soaked in 25% isopropyl alcohol / 10% acetic acid overnight, at room temperature, with gentle rocking, before being stained.

For PAS stain, procedures were modified from Dubray and Bezard (1982) and Hounsell *et al.* (1996). The separating gel was placed in 7.5% acetic acid (enough to cover the gel) and incubated for 30 min, at room temperature, with gentle rocking. The acetic acid was decanted and 100 ml of either 0.2% (w/v) periodic acid in rdH₂O (for detection of 1, 2-diol groups), 1 mM sodium periodate in 0.1 M acetate buffer (pH 5.5) (for the detection of sialic acid) or 15 mM sodium periodate in 0.1 M acetate buffer (pH 5.5) (for the detection of all monosaccharides) were added and the gel left to sit, undisturbed, at 4⁰C, in the dark, for 1 hr. This solution was removed and replaced by

100 ml of Schiff's reagent, and returned to sit at 4⁰C, in the dark, for 1 hr. Schiff's reagent was removed and the gel developed with several changes of 7.5% acetic acid, at room temperature, with gentle rocking until light pink bands developed.

For staining with alcian blue, all volumes used were just enough to cover the gel and solutions were incubated at room temperature, with gentle rocking. The gel was washed with 7.5% acetic acid for 30 min and followed by a rinse with 3% acetic acid for 1 min. Then 1% (w/v) alcian blue in 3% acetic acid was added for 30 min. The stain was removed and the gel destained, first with tap water for 25 min, then with 3% acetic acid overnight.

For staining with thymol-sulfuric acid, the procedure was taken from Racusen (1979). Separating gels were washed in 100 ml of 25% isopropyl alcohol / 10% acetic acid / 0.2% (w/v) thymol for 90 min, at room temperature, with gentle rocking. The solution was removed and replaced by 200 ml of 80% sulfuric acid / 20% ethanol (95% ethanol was used) and rocked at room temperature until light pink bands developed. The gel was made more opaque for photography by the addition of more ethanol to the sulfuric acid-ethanol solution.

2.13.2 Glycosidases

The terminal sugar residue of the glycoprotein was identified using the enzymes neuraminidase (EMD Biosciences Inc.) (removes sialic acids), α -galactosidase (Sigma) (removes alpha-linked galactose), β -galactosidase (Sigma) (removes beta-linked galactose), N-acetyl- β -D-glucosaminidase (Sigma) (removes N-acetyl- β -glucosamine (GlcNAc)) and α -fucosidase (Sigma) (removes alpha-linked fucose). Thirty μ g of Pc ATCC flagellin, Pf ATCC flagellin, Ova, Fet and BSA were combined with 10 μ l of an

enzyme cocktail (made from 2 µl of each enzyme) and the total volume brought to 40 µl with rdH₂O. The proteins were digested for 2 days at 37⁰C. Samples each received 30 µl of SDS-PAGE dissociation solution, were boiled for 5 min and run on an SDS-PAGE gel (section 2.4). Separating gels were stained with Coomassie Brilliant Blue staining solution for 1 hr and destained with Coomassie Blue destain solution overnight.

The N-type carbohydrate linkage to the flagellin peptide backbone was tested for using the endoglycosidase, Peptide: N-Glycosidase F (PNGase F) (New England BioLabs). The reaction conditions were as follows. Twenty µg of protein were denatured with 1x Glycoprotein Denaturing Buffer (5% SDS, 10% β-mercaptoethanol) at 100⁰C for 10 min. A 1/10 volume of both 10x G7 Buffer (0.5 M sodium phosphate, pH 7.5) and 10% NP-40 were added and mixed. Finally, 2 µl (1000 units) of PNGase F were added and the reaction incubated at 37⁰C for 1 hr. The reaction was stopped by the addition of 30 µl of SDS-PAGE dissociation solution. The samples were boiled for 5 min and visualized as described above.

2.14 Phylogenetic analysis

For this study, four different phylogenetic analyses were done: *Pectinatus* flagellin genes (DNA sequence), *Pectinatus* flagellin (deduced aa sequence), *Pectinatus* 16S rRNA gene nucleotide sequence and a comparison of the deduced aa sequences from *Pectinatus* flagellins to other flagellins throughout the bacterial superkingdom. For the analyses within the *Pectinatus* genus, both branch lengths and bootstrap values were calculated, whereas, for the bacterial superkingdom comparison, only a consensus tree with bootstrap values was determined. The analyses within the *Pectinatus* genus also included the flagellin sequence from *R. cecicola* as an outgroup. For each program in

the analysis described below, if a parameter was not specifically stated, the default setting was used. As well, since a 2400 Toshiba Satellite was used for all computations, any time a terminal setting was available (within all Phylogeny Inference Package (PHYLIP) programs), it was changed to IBM PC.

To begin, all sequences were loaded into ClustalX (Thompson *et al.*, 1994; 1997). For *Pectinatus* isolates that were not sequenced completely, N's (for nucleotides) or X's (for aa) were added to fill in the gaps at the very beginning and end of the coding sequences (as compared to the completed sequences). For multiple alignments of DNA sequences, the parameters used were: Gap opening penalty of 30, Gap extension penalty of 6.66, Delay divergent sequences of 30%, DNA transition weight of 0.5, the Use negative matrix was off and the DNA weight matrix was set to IUB. For multiple alignments of protein sequences, the parameters used were: Gap opening penalty of 20, Gap extension penalty of 0.2, Delay divergent sequences of 30%, the Use negative matrix was off and the Protein weight matrix was set to the Gonnet series. Alignments were outputted in Clustal and PHYLIP formats.

Aligned sequences were then manually checked with GeneDoc (Nicholas *et al.*, 1997). The only manual editing done was to realign the *Roseburia* flagellin DNA sequence to maintain the C-terminal domain structure. ClustalX was able to correctly align the domains within the protein sequences, so similar adjustment to the *Roseburia* flagellin protein sequence was unnecessary.

To create the consensus tree and determine bootstrap values, aligned sequences, in PHYLIP format, were first run through SEQBOOT (Felsenstein, 1989) to generate 1,000 variations of the alignment data. Within the parameters, the number of replicates was changed to 1,000. The SEQBOOT outfile was then read into DNADIST /

PROTDIST (Felsenstein, 1989) to determine the distance matrix for each multiple alignment. The Analyze Multiple Data Sets parameter was changed to 1,000. For DNA sequences, the Kimura 2-parameter matrix was used; for protein sequences, the Dayhoff point-accepted mutation (PAM) matrix was used (Kimura, 1980; Dayhoff *et al.*, 1978). The outfile from DNADIST / PROTDIST, containing 1,000 distance matrices, was read into NEIGHBOR (Felsenstein, 1989) to generate 1,000 neighbour-joined phylogenetic trees. Again, the Analyze Multiple Data Sets parameter was changed to 1,000. Both the NEIGHBOR outfile, containing an explanation of each tree generated and the NEIGHBOR treefile, containing instructions to diagram all 1,000 trees generated, were checked. Finally, the NEIGHBOR treefile was read into CONSENSE (Felsenstein, 1989) where one consensus tree, with bootstrap values, was created from analysis of the branching patterns of all 1,000 inputted neighbour-joined trees. The CONSENSE treefile was saved and opened in TreeExplorer to visualize the consensus tree and bootstrap values.

For the three analyses within the *Pectinatus* genus, the analysis was continued to obtain branch lengths for each tree. This was done by running the original data set through DNADIST / PROTDIST, then NEIGHBOUR without multiple data sets, and visualizing the treefile from NEIGHBOR within TreeExplorer. This tree incorporated branch length values into the length of the lines drawn for each branch. Bootstrap values were then manually transferred to the appropriate nodes (Joe Felsenstein, PHYLIP creator, personal communication).

3.0 RESULTS

3.1 Flagellin protein micro-sequencing

Micro-protein sequencing was attempted for five peptides; the intact flagellin protein from Pc ATCC and Pf ATCC, one peptide fragment (29 kDa) from Pc ATCC and two peptide fragments (20 and 29 kDa) from Pf ATCC (Figure 4). The same N-terminal 15 aa sequence was determined from both intact flagellin proteins; namely, ALVVKNMSALNTLN. This agrees with the previously published N-terminal 15 aa for both *Pectinatus* flagellin proteins (Hakalehto *et al.*, 1997). The N-terminal sequence of the peptide fragment from Pc ATCC was found to be NGKLLVDGSM. The two peptides from Pf ATCC resulted in mixed sequencing results and were unsuitable for further analysis.

3.2 Initial sequence of flagellin genes

3.2.1 *Pectinatus cerevisiiphilus*

The PCR primers Flgn2F and Flgn3R were designed from the Pc ATCC aa sequences determined from the N-terminals of the intact flagellin and the 29 kDa flagellin peptide, respectively. These two primers were used to amplify a 419 bp region from the Pc ATCC flagellin gene, using Pc ATCC genomic DNA as the template. This region of flagellin-specific sequence was used as a starting point for Bubble-PCR. All attempts to use Bubble-PCR with gene-specific degenerate primers (FlgnTerminal1 and

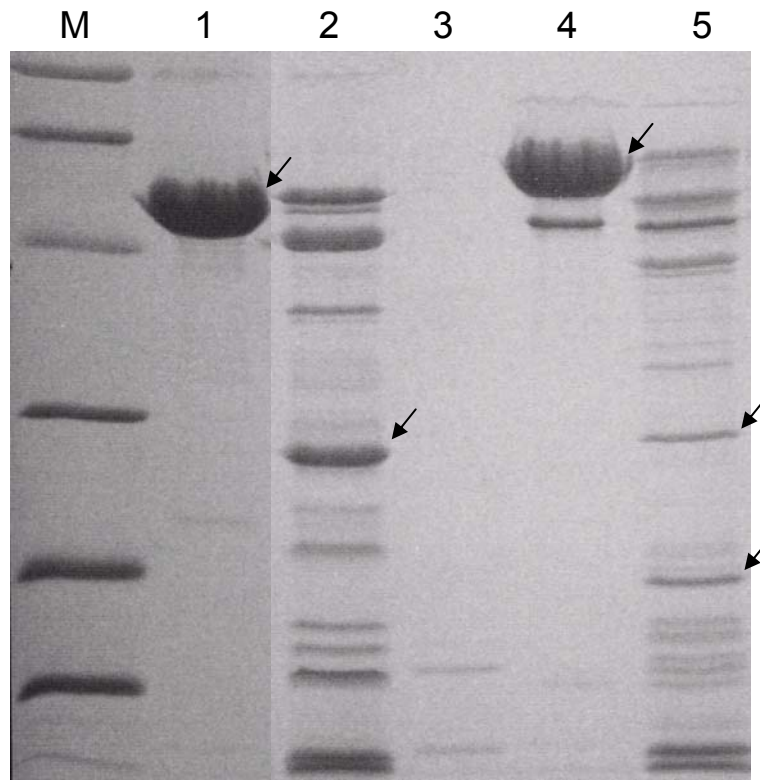


Figure 4 - Alpha-chymotrypsin digest of *Pectinatus* flagellin protein. Lanes are as follows: (M), Marker (97, 66, 45, 31, 21.5 and 14 kDa); (1), Pc ATCC extracted flagellin; (2) Pc ATCC extracted flagellin digested with alpha-chymotrypsin; (3), Alpha-chymotrypsin alone; (4), Pf ATCC extracted flagellin; (5) Pf ATCC extracted flagellin digested with alpha-chymotrypsin. Arrows indicate peptides sent for protein micro-sequencing.

Flgn10F) were unsuccessful, necessitating a specific target DNA sequence to design the gene-specific primer. Through a process of “leap-frogging” over restriction sites, Bubble-PCR was used to amplify and sequence the open reading frame and surrounding DNA sequence of the Pc ATCC flagellin gene. A summary of all PCR primer sets used to complete the sequence of Pc ATCC, including confirmation sequencing, and their amplification products, are found in Table 5. Figure 5 is a schematic of how these PCR products assemble into the completed sequence.

As a note, primer Flgn4R bound to the flagellin gene sequence in two places. This primer was designed to bps 373-394 to be used in Bubble-PCR to sequence over the promoter region of the gene (product A in Table 5 and Figure 5). However, by chance, 13 of the 22 bp in primer Flgn4R are identical to the downstream sequence at 1006-1027 bp. When PCR was carried out at 40⁰C annealing temperature with Pc ATCC genomic DNA and primers Flgn5F and Flgn4R, a 435 bp PCR product was generated (product D in Table 5 and Figure 5). In Figure 5, the partial binding of primer Flgn4R to Pc ATCC flagellin at 1006-1027 bp is represented by a small box and a tailing line.

To confirm the flagellin sequence independently of Bubble-PCR, Pc ATCC primers from the last third of the sequence were paired with the primers Cterm1 and Cterm2. The Cterm primers are flagellin-specific primers based on the consensus sequence of many known flagellin genes. The products of the primers Cterm1 and Cterm2 with Pc ATCC primers were sequenced to confirm the gene identity as flagellin. Nearly all of the Pc DSM and Pc VTT flagellin genes were amplified and sequenced from one round of PCR using the PectCloneF and PectCloneR primer set (Table 4).

Table 5 – PCR product composition of *Pectinatus cerevisiiphilus* ATCC 29359^T.

PCR product^a	Forward Primer	Reverse Primer	Length of Sequence (bp)^b
A	BubblePrimer (RsaI)	Flgn4R	394
B	Flgn2F	Flgn3R	419
C	Flgn5F	BubblePrimer (HaeIII)	653
D	Flgn5F	Flgn4R	435
E	Flgn5F	BubblePrimer (RsaI)	340
F	Flgn6F	Cterm1	423
G	Flgn7F	Cterm1	362
H	Flgn7F	BubblePrimer (RsaI)	138
I	Flgn8F	BubblePrimer (EcoRV)	309
J	Flgn8F	Cterm2	263
K	Flgn8F	BubblePrimer (HaeIII)	180

^a PCR product letters refer to Figure 5.

^b Lengths with a BubblePrimer restriction site count from the flagellin primer (including the primer length in the sequence) up to the middle of the restriction site (does not include the Bubble-linker or BubblePrimer length). Lengths with two non-Bubble primers include the length of both primers in the sequence length.

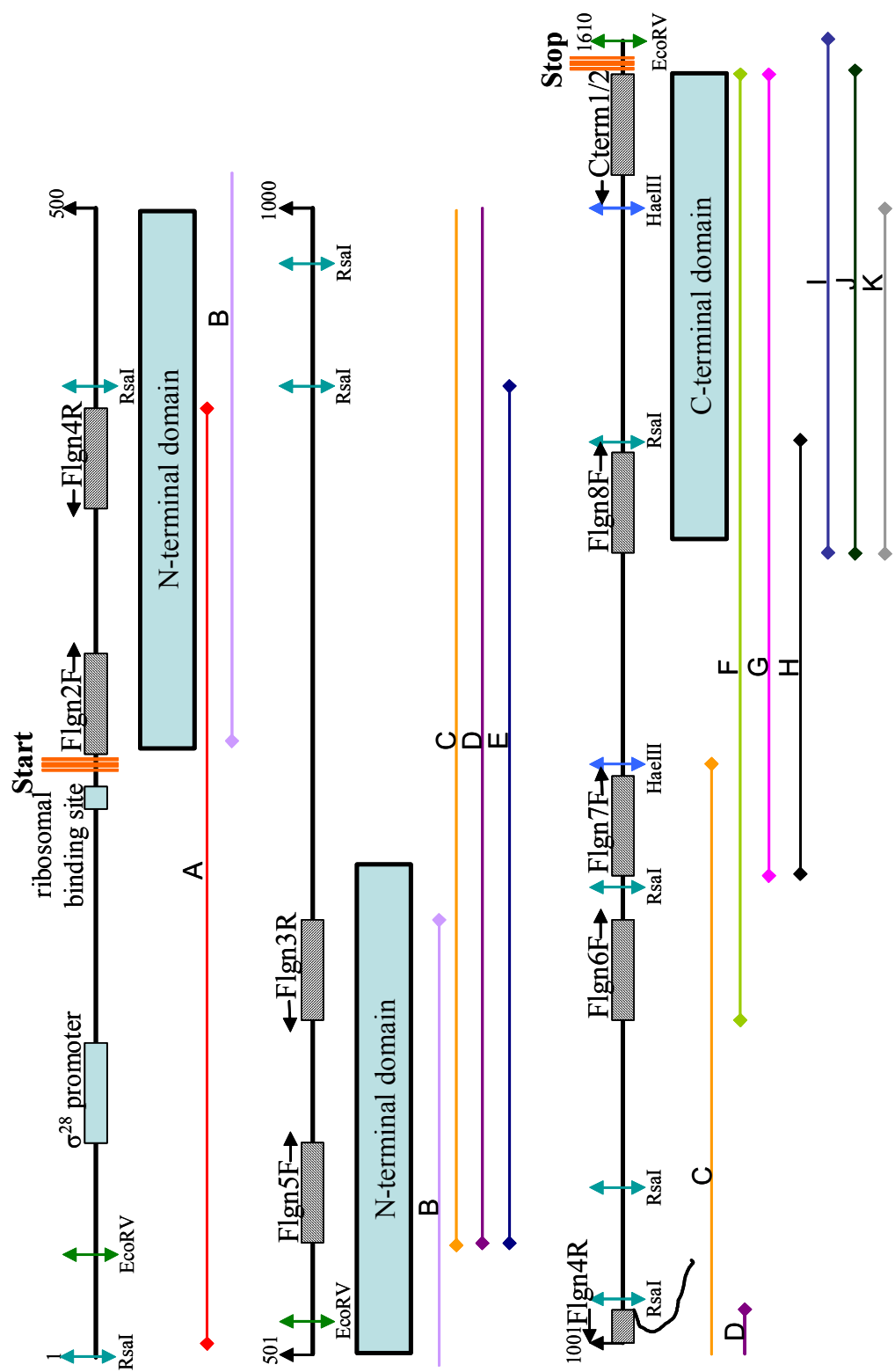


Figure 5 – Schematic of the PCR assembly of *Pectinatus cerevisiophilus* ATCC 29359^T flagellin gene. PCR product labels correspond to Table 5. The partial binding of primer Flgn4R to the sequence at 1006-1027 bp is represented by a small box and a tailing line.

3.2.2 *Pectinatus frisingensis*

A 257 bp sequence from the Pf DSM flagellin gene was obtained by PCR with the primers Flgn8F and Cterm1, using Pf DSM genomic DNA. Although the primer Flgn8F was designed from Pc ATCC flagellin sequence, there is enough sequence similarity in this region of *Pectinatus* flagellin genes to allow for primer binding. A combination of PCR and Bubble-PCR was used to sequence the remainder of the gene. Bubble-PCR was not used to sequence the entire gene because primer Flgn13R did not work with the Bubble system. It remains unclear why this primer is incompatible with Bubble-PCR. Flgn13R did generate the expected Pf DSM flagellin PCR product with Flgn9F (a specific PCR primer designed to the N-terminal region of Pc ATCC), allowing Bubble-PCR to be resumed upstream. The PCR product from Flgn8F and Cterm2 was generated and sequenced as an additional confirmation. Table 6 contains the summary of PCR primers and products used to sequence the Pf DSM flagellin gene and Figure 6 diagrams the PCR product assembly. Essentially all of Pf ATCC and Pf VTT flagellin genes were amplified and sequenced from one round of PCR using the PectCloneF and PectCloneR primer set.

3.3 Expression and confirmation of flagellin genes

The flagellin genes identified from Pc ATCC and Pf DSM were confirmed by immunoblotting experiments (Figure 7). The theoretical sizes of the GST-flagellin fusion proteins are 73 kDa (the predicted size of the flagellin proteins from the translated sequence is 47 kDa plus the GST tag size of 26 kDa). The fusion proteins were detected in pGEX-Pc and pGEX-Pf transfected cells at approximately 68 kDa with both anti-flagellin and anti-GST antibody (lanes 2 and 3, respectively). This is within 10% of the

Table 6 – PCR product composition of *Pectinatus frisingensis* DSM 20465.

PCR product^a	Forward Primer	Reverse Primer	Length of Sequence (bp)^b
AA	BubblePrimer (RsaI)	Flgn14R	314
BB	Flgn9F	Flgn13R	369
CC	BubblePrimer (EcoRV)	Flgn12R	924
DD	Flgn8F	Cterm1	257
EE	Flgn8F	Cterm2	263
FF	Flgn11F	BubblePrimer (RsaI)	311

^a PCR product letters refer to Figure 6.

^b Lengths with a BubblePrimer restriction site count from the flagellin primer (including the primer length in the sequence) up to the middle of the restriction site (does not include the Bubble-linker or BubblePrimer length). Lengths with two non-Bubble primers include the length of both primers in the sequence length.

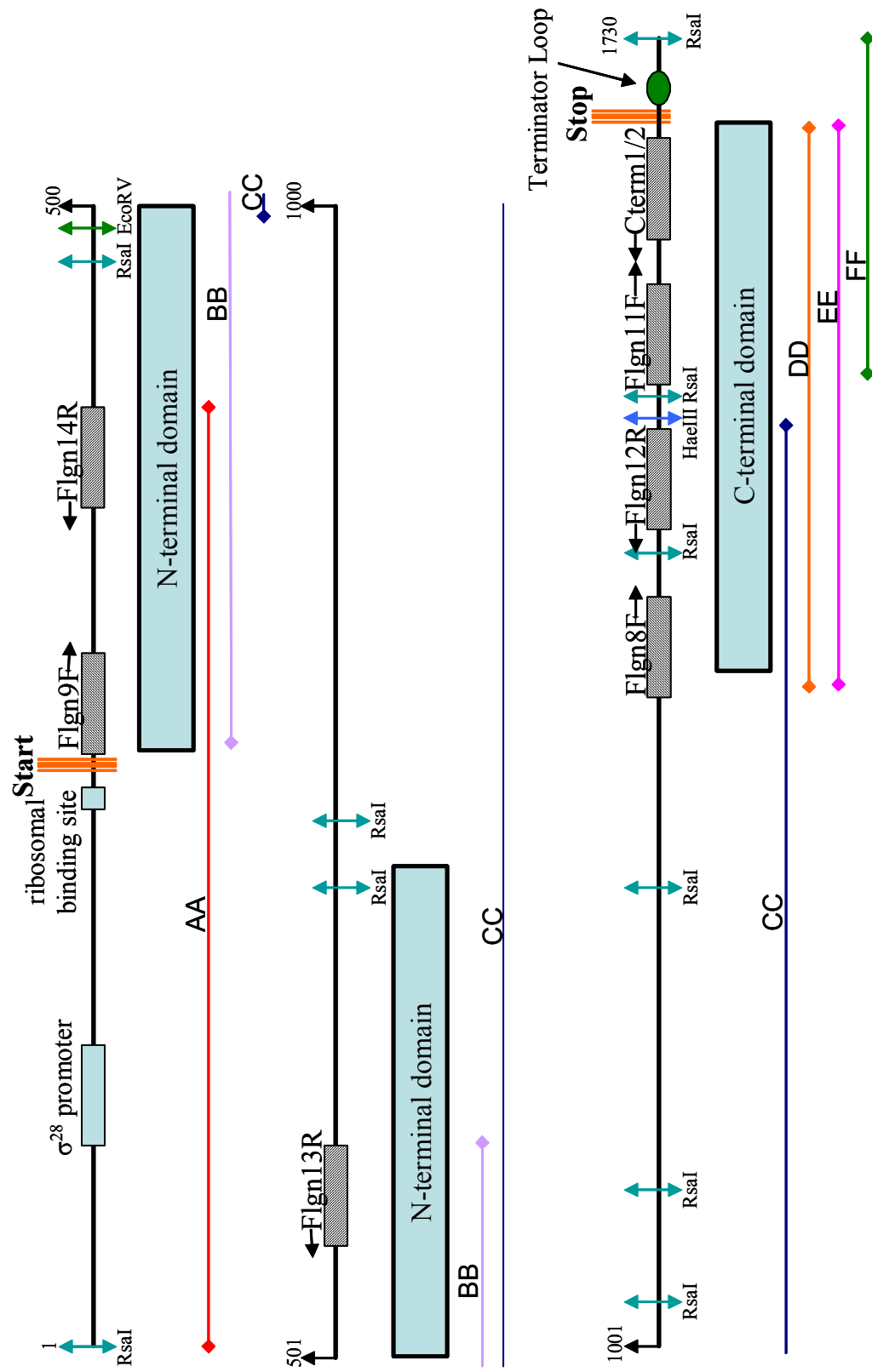


Figure 6 – Schematic of the PCR assembly of *Pectinatus frisingensis* DSM 20466 flagellin gene. PCR product labels correspond to Table 6.

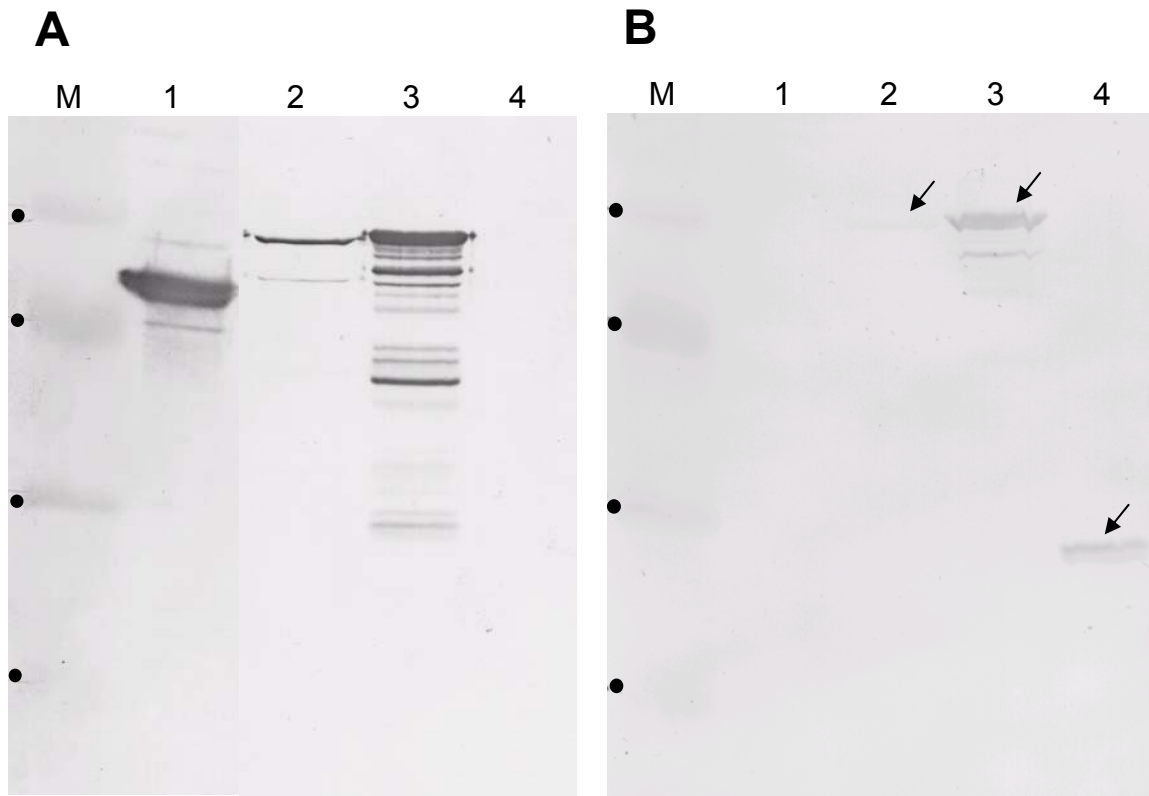


Figure 7 – Immunoblots of GST-flagellin fusion proteins. The blots were developed with (A) anti-*Pectinatus* flagellin antibody (1.25 hour post-induction) or (B) anti-GST antibody (1.5 hour post-induction). Lanes were as follows: (M) marker (68, 43, 29 and 18 kDa); (1) pooled, purified Pc DSM and Pc VTT flagella; extracts of (2) pGEX-Pc, (3) pGEX-Pf, and (4) pGEX control transfected *E. coli*. Arrows indicate faint bands.

predicted molecular weight. As well, the ladder-like banding pattern of the fusion proteins indicate that they are being degraded within the host *E. coli* cells. Finally, GST protein was detected from the pGEX vector alone at the expected 26 kDa with anti-GST antibody (blot B, lane 4).

3.4 *Pectinatus* flagellin characteristics

The completed flagellin sequences from all six *Pectinatus* isolates studied were aligned using ClustalX and visualized using GeneDoc. The DNA sequence alignment of the flagellin genes is shown in Figure 8 and the flagellin protein alignment is shown in Figure 9. All DNA and protein sequence numbering is in reference to Figures 8 and 9, respectively. The four flagellin genes that were sequenced from the PectClone primer set, namely, Pc DSM, Pc VTT, Pf ATCC and Pf VTT, do not contain the complete open reading frame of the flagellin gene. This is because the primer set bound to the N- and C-terminal portions of the genes. To sequence the approximately 1300 bp of the gene with the PectClone primer set, sequencing from both the forward and reverse primer was necessary. Since the sequencing method does not generate the primer sequence or sequence immediately downstream of the primer, these regions of the flagellin gene were not determined. The open reading frame of the complete flagellin genes begins at 250 bp and ends at 1596 bp (Figure 8). The Pc ATCC flagellin coding region is 1338 bp (446 aa) while the Pf DSM flagellin coding region is 1344 bp (448 aa). The sequence gap between the two species always occurs at the same position, 916-921 bp, corresponding to amino acids 221 and 222. From both alignments, the conservation of the N- and C-terminal domains is clearly evident.

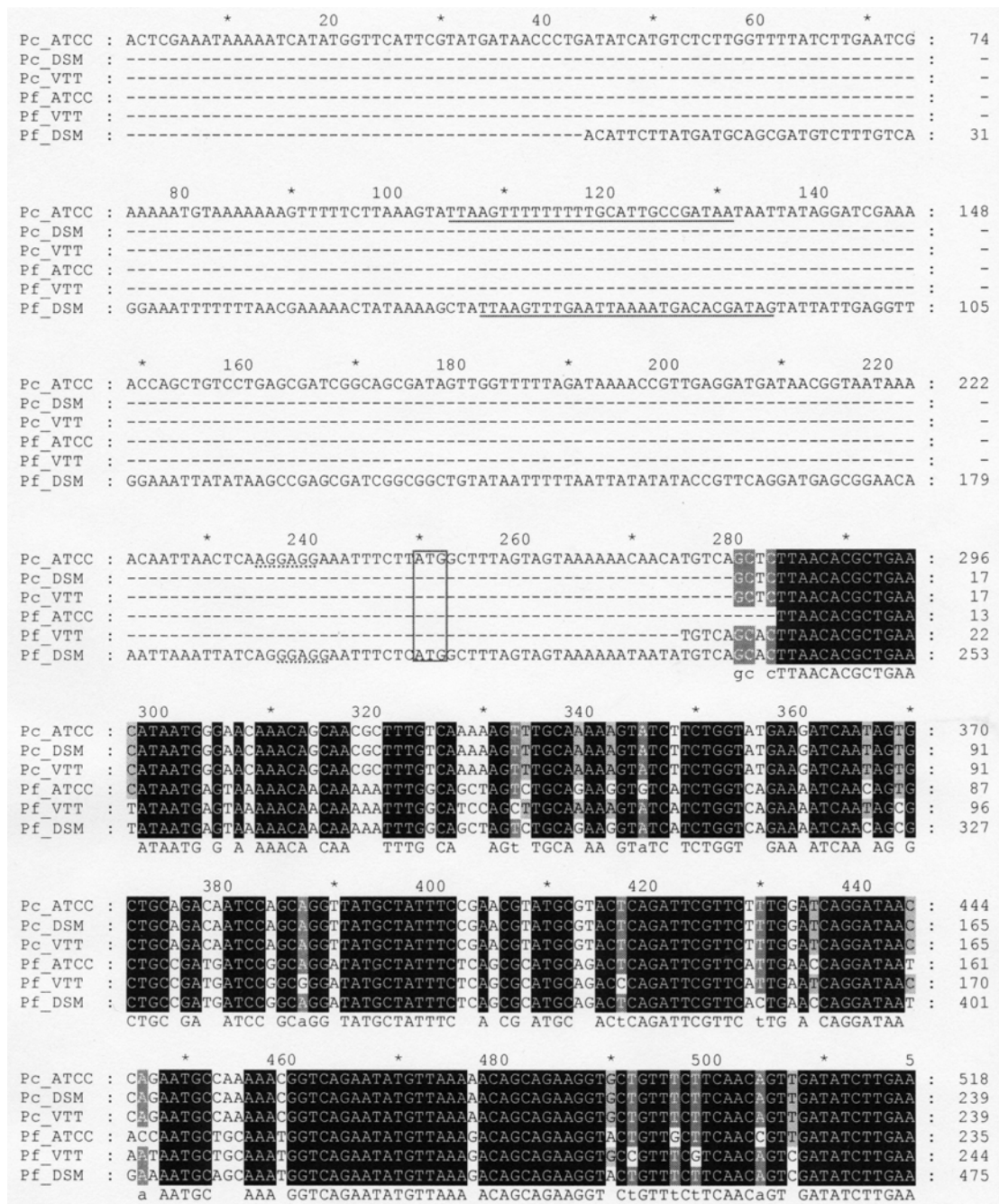


Figure 8 – ClustalX alignment of *Pectinatus* flagellin gene sequences. Alignment parameters are stated in Materials and Methods, and final alignment was viewed in GeneDoc. The consensus sequence is displayed under the alignment where uppercase letters represent 6 of 6 identical nucleotides and lowercase letters represent 5 of 6 identical nucleotides. Shading is as follows: 6 of 6 nucleotides identical - white text on black background; 5 of 6 nucleotides identical - white text on grey background; 4 of 6 nucleotides identical - black text on grey background. Dashes (-) in the sequences represent gaps, solid rectangle indicates start codon, dashed rectangle indicates stop codon, solid line indicates putative sigma 28 promoter sites, dashed line indicates ribosomal binding sites and arrows indicate terminating stem-loop.

```

20          *          540          *          560          *          580          *
Pc_ATCC : AACCTTGAAGGAAAAAGTTCTTGATGCGCCCAATGATACCAATACAAAGCTTTGACCGCAATACATTAGGATT : 592
Pc_DSM : AACCTTGAAGGAAAAAGTTCTTGATGCGCCCAATGATACCAATACAAAGCTTTGACCGCAATACATTAGGATT : 313
Pc_VTT : AACCTTGAAGGAAAAAGTTCTTGATGCGCCCAATGATACCAATACAAAGCTTTGACCGCAATACATTAGGATT : 313
Pf_ATCC : AACCTTGAAGGAAAAAGTTCTTGATGCGCCCAATGATACCAATACAAAGCTTTGACCGCAATACATTAGGATT : 309
Pf_VTT : AACCTTGAAGGAAAAAGTTCTTGATGCGCCCAATGATACCAATACAAAGCTTTGACCGCAATACATTAGGATT : 318
Pf_DSM : AACCTTGAAGGAAAAAGTTCTTGATGCGCCCAATGATACCAATACAAAGCTTTGACCGCAATACATTAGGATT : 549
AAC TTGAAGGAAAAAGTTCTTGATGCGCCCAATGATACCAATAC AC T TGA CGCAATAC ATTAGGATT

600          *          620          *          640          *          660
Pc_ATCC : ATATGATCAGGCTATAGACAGATCAATGACAATGCCAATGCTTTTATAATGGTAAATTTCTTGTTGAGGGT : 666
Pc_DSM : ATATGATCAGGCTATAGACAGATCAATGACAATGCCAATGCTTTTATAATGGTAAATTTCTTGTTGAGGGT : 387
Pc_VTT : ATATGATCAGGCTATAGACAGATCAATGACAATGCCAATGCTTTTATAATGGTAAATTTCTTGTTGAGGGT : 387
Pf_ATCC : ACTTGGATCAGGCTATAGACAGATCAATGACAATGCCAATGCTTTTATAATGGTAAATTTCTTGTTGAGGGT : 383
Pf_VTT : ACTTGGATCAGGCTATAGACAGATCAATGACAATGCCAATGCTTTTATAATGGTAAATTTCTTGTTGAGGGT : 392
Pf_DSM : ACTTGGATCAGGCTATAGACAGATCAATGACAATGCCAATGCTTTTATAATGGTAAATTTCTTGTTGAGGGT : 623
A TG ATCAGGc T GacCAGATcAA GA AATGC Aatg CttATAATgg AA ct CttGTTGATGGT

680          *          700          *          720          *          740
Pc_ATCC : TCTATGAATAAAGCAACTGCTGGTGGTACCCTAAATATACTAATGAAAGTTTAGGCTACCGACTGATTTTGG : 740
Pc_DSM : TCTATGAATAAAGCAACTGCTGGTGGTACCCTAAATATACTAATGAAAGTTTAGGCTACCGACTGATTTTGG : 461
Pc_VTT : TCTATGAATAAAGCAACTGCTGGTGGTACCCTAAATATACTAATGAAAGTTTAGGCTACCGACTGATTTTGG : 461
Pf_ATCC : TCTATGAATAAAGCAACTGCTGGTGGTACCCTAAATATACTAATGAAAGTTTAGGCTACCGACTGATTTTGG : 457
Pf_VTT : TCTATGAATAAAGCAACTGCTGGTGGTACCCTAAATATACTAATGAAAGTTTAGGCTACCGACTGATTTTGG : 466
Pf_DSM : TCTATGAATAAAGCAACTGCTGGTGGTACCCTAAATATACTAATGAAAGTTTAGGCTACCGACTGATTTTGG : 697
TC ATGAAT Ag tG t C AC GT C TAt C AATG AAGTTtag aC g AC TTT

760          *          780          *          800          *
Pc_ATCC : TTCTGCATTAACTGATCTAGTTGATAGAAATGGTAAATCACTCAATATTTAGACAAAGATAGTTTCTATTT : 814
Pc_DSM : TTCTGCATTAACTGATCTAGTTGATAGAAATGGTAAATCACTCAATATTTAGACAAAGATAGTTTCTATTT : 535
Pc_VTT : TTCTGCATTAACTGATCTAGTTGATAGAAATGGTAAATCACTCAATATTTAGACAAAGATAGTTTCTATTT : 535
Pf_ATCC : TTCTGCATTAACTGATCTAGTTGATAGAAATGGTAAATCACTCAATATTTAGACAAAGATAGTTTCTATTT : 531
Pf_VTT : TTCTGCATTAACTGATCTAGTTGATAGAAATGGTAAATCACTCAATATTTAGACAAAGATAGTTTCTATTT : 540
Pf_DSM : TTCTGCATTAACTGATCTAGTTGATAGAAATGGTAAATCACTCAATATTTAGACAAAGATAGTTTCTATTT : 771
T C aTT AC G TA GAT G AATGG aaTtC CT AATATT T c a GAtac GTT C T T

820          *          840          *          860          *          880
Pc_ATCC : CTTATGTTGTCAACGGCAAGGCTATGACAACAGCTTTTCTCTGGTTCAAGTAAATTTCTGATGTTTGTGGC : 888
Pc_DSM : CTTATGTTGTCAACGGCAAGGCTATGACAACAGCTTTTCTCTGGTTCAAGTAAATTTCTGATGTTTGTGGC : 609
Pc_VTT : CTTATGTTGTCAACGGCAAGGCTATGACAACAGCTTTTCTCTGGTTCAAGTAAATTTCTGATGTTTGTGGC : 609
Pf_ATCC : CTTATGTTGTCAACGGCAAGGCTATGACAACAGCTTTTCTCTGGTTCAAGTAAATTTCTGATGTTTGTGGC : 605
Pf_VTT : CTTATGTTGTCAACGGCAAGGCTATGACAACAGCTTTTCTCTGGTTCAAGTAAATTTCTGATGTTTGTGGC : 614
Pf_DSM : CTTATGTTGTCAACGGCAAGGCTATGACAACAGCTTTTCTCTGGTTCAAGTAAATTTCTGATGTTTGTGGC : 845
CTTATGT AACGG AA a TATGACAAC AG Ttt GttGg agTA Tta tgat T TT

900          *          920          *          940          *          960
Pc_ATCC : AGTGTTCACAAGTGCTACAGGCTACAGT-----GATTTTCTAATAGTGTCTTTTGGCATACTTTTGCCCA : 956
Pc_DSM : AGTGTTCACAAGTGCTACAGGCTACAGT-----GATTTTCTAATAGTGTCTTTTGGCATACTTTTGCCCA : 677
Pc_VTT : AGTGTTCACAAGTGCTACAGGCTACAGT-----GATTTTCTAATAGTGTCTTTTGGCATACTTTTGCCCA : 677
Pf_ATCC : AAGGCTAATCAGGCTACAGGCGGCACTATAACAGCAAGGGGCAATAGTGTCTTTACTTAAACCGGGGGAAGTGG : 679
Pf_VTT : AAGGCTAATCAGGCTACAGGCGGCACTATAACAGCAAGGGGCAATAGTGTCTTTACTTAAACCGGGGGAAGTGG : 688
Pf_DSM : AAGGCTAATCAGGCTACAGGCGGCACTATAACAGCAAGGGGCAATAGTGTCTTTACTTAAACCGGGGGAAGTGG : 919
a G T Gct c GG ca T G T G taAT gtgttt TT a c t g

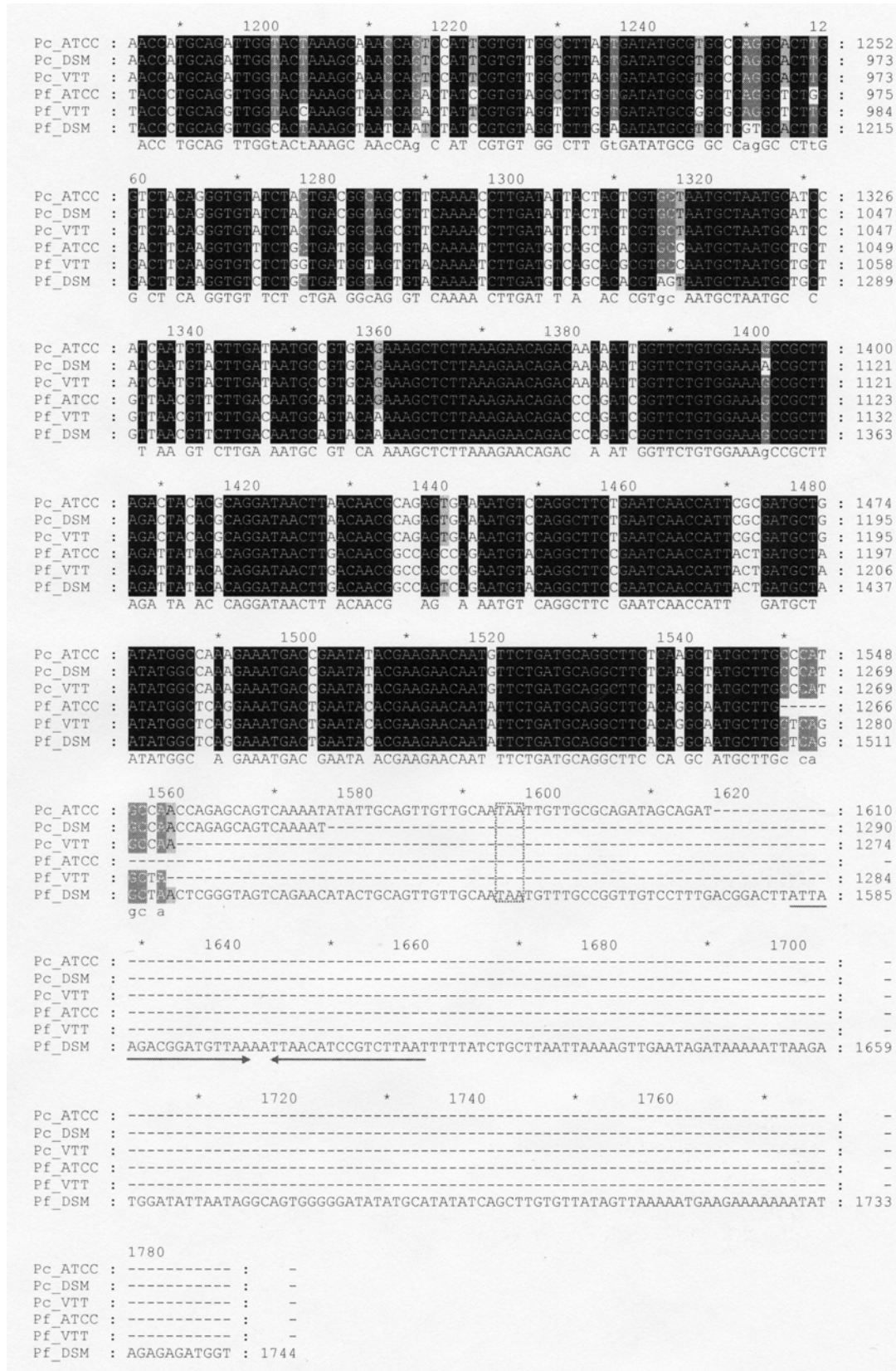
980          *          1000          *          1020          *
Pc_ATCC : CAGCCAAATGGTACAACTGCTGAGGTTCTAATGATGCACTGTCAGACGGTACATCAGTATTAGCAATTTCTT : 1030
Pc_DSM : CAGCCAAATGGTACAACTGCTGAGGTTCTAATGATGCACTGTCAGACGGTACATCAGTATTAGCAATTTCTT : 751
Pc_VTT : CAGCCAAATGGTACAACTGCTGAGGTTCTAATGATGCACTGTCAGACGGTACATCAGTATTAGCAATTTCTT : 751
Pf_ATCC : TTCTACATTTGGTAAAGATGCCAGGCAAGTGGCAATGCAACAGCAGATGGTCTCTGTTTGAAGTGTACTT : 753
Pf_VTT : TTCTACATTTGGTAAAGATGCCAGGCAAGTGGCAATGCAACAGCAGATGGTCTCTGTTTGAAGTGTACTT : 762
Pf_DSM : AAGCTTTTATGGGTTTAAATGCACAGGTTGATCCGATATAGGCGAGATGGTCAAGTGTATGGGAGTTGCTT : 993
aAtTGG aA TGC AaGgt T a atGcAtAc GCAGA Ggt C tc Gt TT Ct TT CTT

1040          *          1060          *          1080          *          1100
Pc_ATCC : CAGTGCATCAGGCTTGGGCGCAATATTTGGTTTACTATCAGCATCCTGGTACTGACGTAATGTCAA : 1104
Pc_DSM : CAGTGCATCAGGCTTGGGCGCAATATTTGGTTTACTATCAGCATCCTGGTACTGACGTAATGTCAA : 825
Pc_VTT : CAGTGCATCAGGCTTGGGCGCAATATTTGGTTTACTATCAGCATCCTGGTACTGACGTAATGTCAA : 825
Pf_ATCC : CAGTGCATCAGGCTTGGGCGCAATATTTGGTTTACTATCAGCATCCTGGTACTGACGTAATGTCAA : 827
Pf_VTT : CAGTGCATCAGGCTTGGGCGCAATATTTGGTTTACTATCAGCATCCTGGTACTGACGTAATGTCAA : 836
Pf_DSM : CAGTGCATCAGGCTTGGGCGCAATATTTGGTTTACTATCAGCATCCTGGTACTGACGTAATGTCAA : 1067
C A T GG gTt G GG AaTat gC GG TTTAC ATCAGcAt aCTGGTAC GA GGTAAATgT AAa

1120          *          1140          *          1160          *          1180
Pc_ATCC : AATATGTAATTTTAAAGCTTGATGCGTGGAGTGAACAGATAGCSCCAGGATAAATCAAATGATAATGTGT : 1178
Pc_DSM : AATATGTAATTTTAAAGCTTGATGCGTGGAGTGAACAGATAGCSCCAGGATAAATCAAATGATAATGTGT : 899
Pc_VTT : AATATGTAATTTTAAAGCTTGATGCGTGGAGTGAACAGATAGCSCCAGGATAAATCAAATGATAATGTGT : 899
Pf_ATCC : AATATGTAATTTTAAAGCTTGATGCGTGGAGTGAACAGATAGCSCCAGGATAAATCAAATGATAATGTGT : 901
Pf_VTT : AATATGTAATTTTAAAGCTTGATGCGTGGAGTGAACAGATAGCSCCAGGATAAATCAAATGATAATGTGT : 910
Pf_DSM : AATATGTAATTTTAAAGCTTGATGCGTGGAGTGAACAGATAGCSCCAGGATAAATCAAATGATAATGTGT : 1141
AaATC GT aAT cTaaGCTTGATGC TGG tGAACAGATCA GC CAGGATAAATCAAATGATAAT C T

```

Figure 8 - ClustalX alignment of *Pectinatus* flagellin gene sequences (continued).



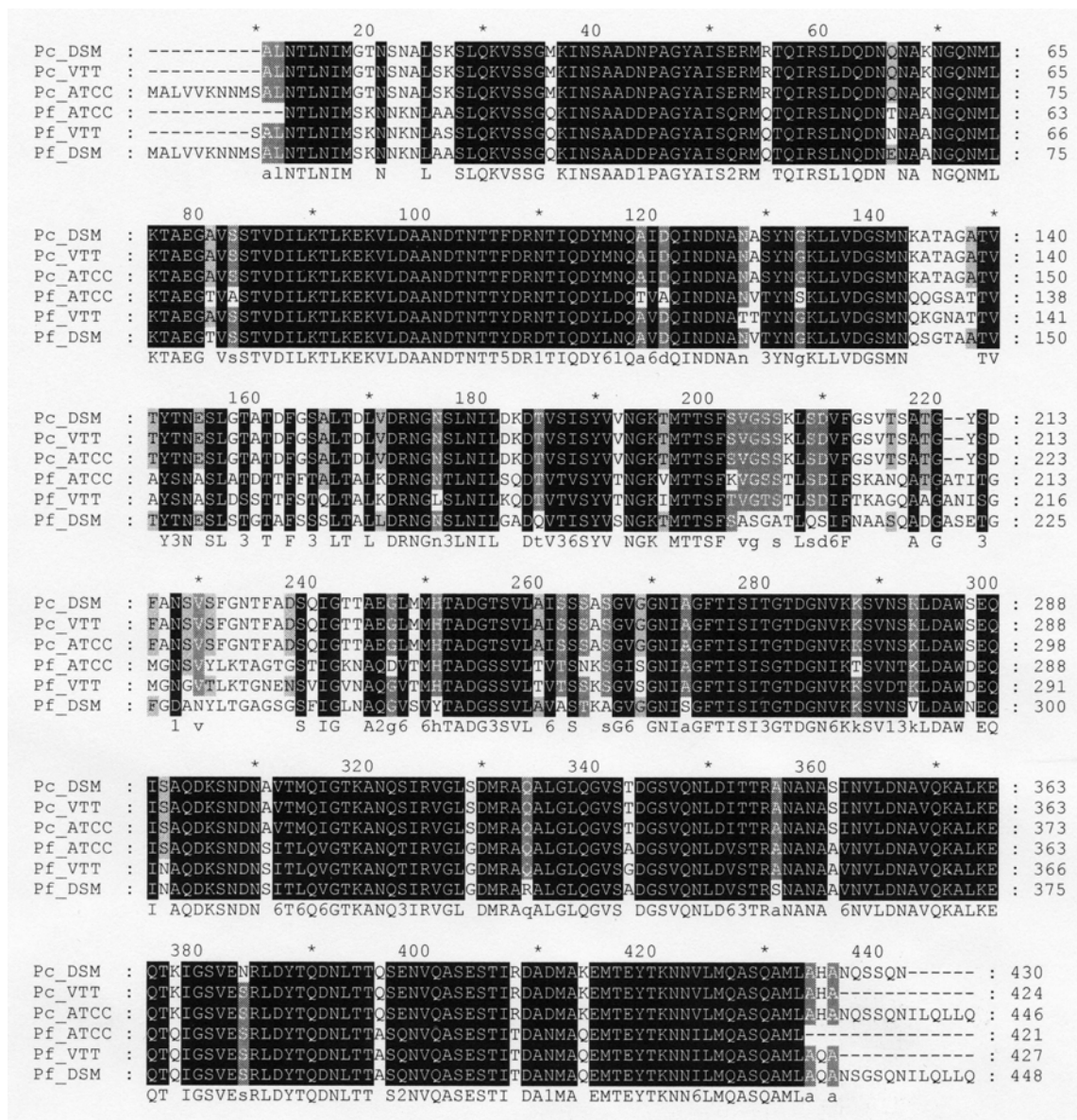


Figure 9 – ClustalX alignment of *Pectinatus* flagellin proteins. Alignment parameters are stated in Materials and Methods and final alignment was viewed in GeneDoc. The consensus sequence is displayed under the alignment where uppercase letters represent 6 of 6 identical aa, lowercase letters represent 5 of 6 identical aa and numbers represent conserved substitutions (1= D + N; 2 = E + Q; 3 = S + T; 5 = F + Y; 6 = I + V). Shading is as follows: 6 of 6 aa identical - white text on black background; 5 of 6 aa identical - white text on grey background; 4 of 6 aa identical - black text on grey background. Dashes (-) in the sequences represent gaps.

Outside of the open reading frames, DNA sequences from Pc ATCC and Pf DSM reveal the expected regulator elements. Putative sigma 28 promoter sites are found approximately 130 bp upstream of the start codon, at 105-131 bp in Pc ATCC and 108-135 bp in Pf DSM from Figure 8. A comparison of these regions to known sigma 28 promoter sites is shown in Figure 10. At approximately -10 bp from the start codon, ribosomal binding sites are evident in both sequences. The sequence AGGAGG begins at 236 bp in Pc ATCC and the sequence GGAGG begins at 237 bp in Pf DSM. This is most of the traditional Shine-Dalgarno ribosomal binding site of AGGAGGU (Snyder and Champness, 1997). Finally, a terminating loop structure can be found in the Pf DSM sequence from 1625 bp to 1660 bp, creating a stem of 17 bp and a loop of 2 bp. A comparable loop structure was not determined for the flagellin gene of Pc ATCC because the final Bubble-PCR product did not extend downstream of the stop codon far enough to extend over this region.

3.5 Glycosylation analysis via staining and glycosidases

Two methods were used to test for glycosylation of *Pectinatus* flagellin proteins; carbohydrate-specific stains and glycosidases. A summary of the results are given in Table 7. Staining results are shown in Figure 11. PAS stain was used to detect the presence of 1, 2-diol groups and sialic acid (Dubray and Bezard, 1982; Hounsell *et al.*, 1996). Five of the six flagellin proteins (Pf ATCC being the exception) were positive for 1, 2-diol groups. Only the type stain from each species was tested for sialic acid and both were negative. An attempt was made to modify the PAS stain to detect all monosaccharides, however, this resulted in over-oxidation of the protein backbone and all proteins were visualized. The second stain used was the thymol-sulfuric acid stain.

Pc ATCC	TTAAGTTT	-11-	GCC CGATAA	(This study)
Pf DSM	TTAAGTTT	-12-	CAC CGATAG	(This study)
<i>E. coli</i> and <i>S. typhimurium</i>				
original consensus	TAAA	-15-	GCC CGATAA	(Helmann et al, 1987)
updated consensus	TAAAGTTT	-11-	GCC CGATAA	(Ide et al, 1999)
<i>Vibrio parahaemolyticus</i>				
consensus	TAAA	-15-	GCC CGTTAA	(McCater, 1995)
<i>P. aeruginosa</i>				
fliC	TAAAG	-14-	GCC CGATAA	(Spangenberg et al, 1996)
<i>B. subtilis</i>				
consensus	CTAAA	-16-	CC CGATAT	(Gilman et al, 1981)
<i>C. jejuni</i>				
flaA	ATAAA	-15-	AAC CGATAT	(Nuijten et al, 1990)
<i>Treponema phagedenis</i>				
flaB2	CTTAA	-15-	TCC CGATAC	(Limberger et al, 1992)

Figure 10 – Sigma 28 flagellin promoters. Nucleotides in bold are identical in all of the sequences listed.

Table 7 – Summary of carbohydrate detection.

Proteins	Stains			Enzymes (glycosidases)	
	Thymol -Sulfuric Acid	Periodic Acid – Schiff's		Endoglycosidase	Exoglycosidase
		1,2-Diols	Sialic Acid	Neuraminidase	PNGase F
Pc ATCC	+	+	-	-	-
Pc DSM	+	+	nd ^a	nd	nd
Pc VTT	+	+	nd	nd	nd
Pf ATCC	-	-	-	-	-
Pf DSM	+ ^b	+	nd	nd	nd
Pf VTT	+ ^b	+	nd	nd	nd
Ova	-	+	-	-	+
Fet	-	+	+	+	+
HRP	+	+	nd	nd	nd
BSA	-	-	-	-	-

^a Not done.

^b Very weak – 1/10th the intensity of other positive bands.

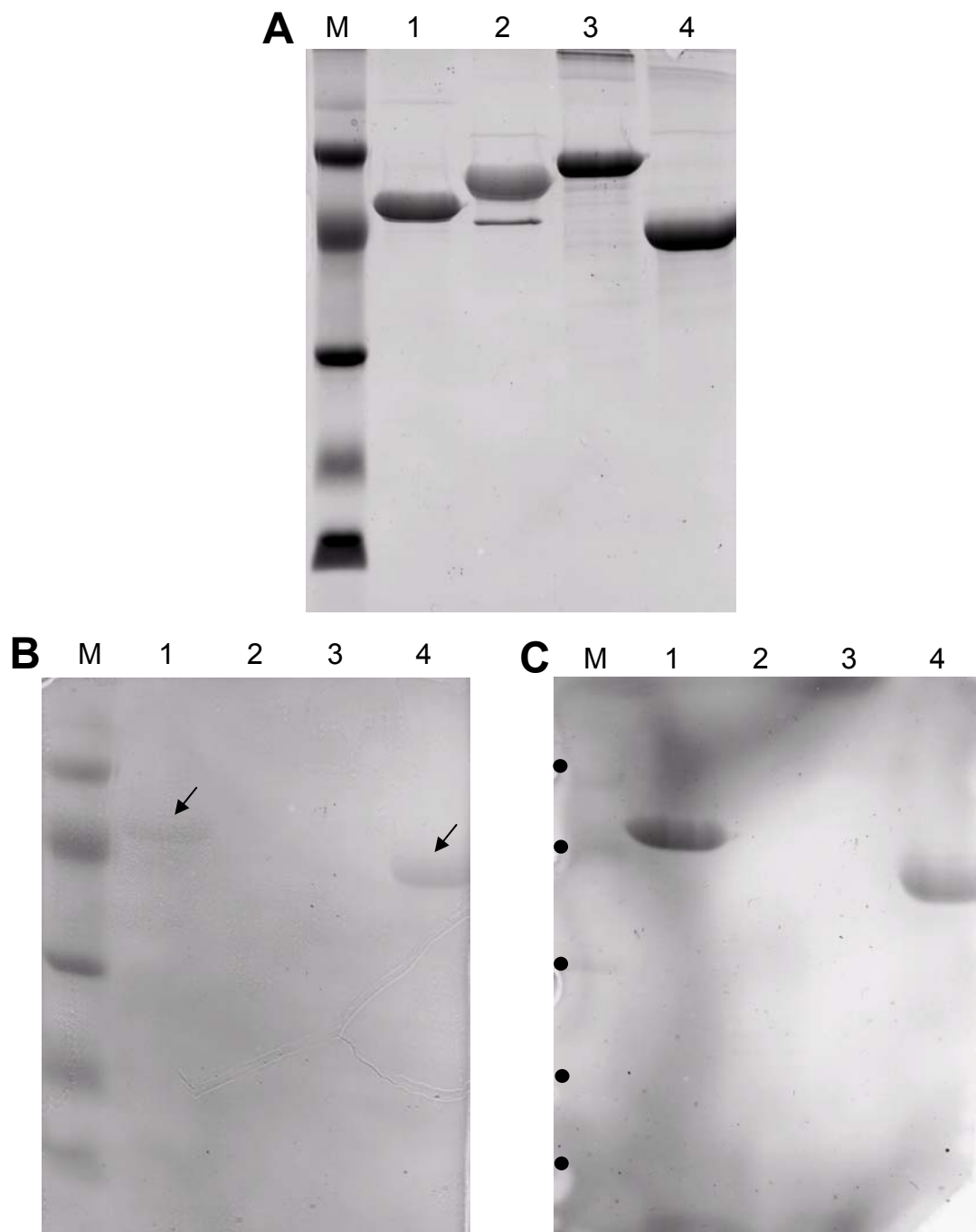


Figure 11 – Carbohydrate staining results. Gels were developed with (A) Coomassie Blue stain, (B) PAS stain and (C) thymol-sulfuric acid stain. Lanes were as follows: (M), marker (68, 43, 29, 18 and 13 kDa); (1), Pc ATCC flagellin protein; (2), Pf ATCC flagellin protein; (3), BSA protein; (4), Ova protein (for A and B) or HRP protein (for C). Arrows indicate faint bands.

This stain reacts with carbohydrate residues that contain a C-2 hydroxyl group (Hounsell *et al.*, 1996 and references within). Except for 2-amino sugars, this reaction is sensitive and general for pentoses, hexoses, uronic acids and all of their polymers (Racusen, 1979). Under these conditions, all three Pc flagellin proteins stained brightly, Pf DSM and Pf VTT stained weakly and Pf ATCC did not stain. A third carbohydrate-specific stain, alcian blue, was also tried. Under the conditions used in this study, alcian blue was found to be incompatible with SDS-PAGE separated proteins and all proteins present were stained.

Two different classes of glycosidases were used to detect carbohydrate presence on *Pectinatus* flagellin proteins; endoglycosidases and an exoglycosidase. Endoglycosidases remove the terminal sugar from a carbohydrate chain. In this study, the endoglycosidases neuraminidase (removes sialic acid), α - and β -galactosidase (removes both forms of galactose), α -fucosidase (removes fucose) and N-acetyl- β -D-glucoaminidase (removes GlcNAc) were tested. All enzymes had no effect on the flagellin type stain proteins. Only neuraminidase had a positive control (fet). Since the other enzymes did not have a positive control to ensure the enzyme was functioning properly, it is premature to conclusively state that *Pectinatus* flagellin do not contain any of these other sugars as their terminal residue. Exoglycosidases remove carbohydrate chains from proteins at the carbohydrate-protein linkage. The exoglycosidase used was PNGase F, which hydrolyzes nearly all types of N-linked glycan chains from glycoproteins. The only N-linkage known to be resistant to PNGase F is a carbohydrate containing a core α 1-3 fucose. Pc ATCC and Pf ATCC flagellin proteins were resistant to PNGase F cleavage and, as such, appear not to be N-linked. This suggests that the

glycan linkage for *Pectinatus* flagellin proteins is O-linked. Unfortunately, there are no exoglycosidases available to remove all O-linked carbohydrates.

3.6 Taxonomic analysis of *Pectinatus* via phylogenetics

Taxonomic information for *Pectinatus* was collected and assessed from four phylogenetic analyses. The first comparison was done at the superkingdom level by analyzing the placement of *Pectinatus* flagellin proteins within a framework of 30 known flagellin proteins (Figure 12). The sources of the known flagellin sequences are listed in Table 3. The phylogenetic tree shows that *Pectinatus* isolates cluster together with the phylum *Firmicutes*. The second and third comparisons analyze the relatedness of the *Pectinatus* flagellin at the DNA and protein levels (Figures 13A and 13B, respectively). Both phylogenetic trees show that *P. cerevisiiphilus* isolates are more closely related than the isolates from *P. frisingensis*. Finally, the fourth comparison was done with the *Pectinatus* 16S rRNA gene as a reference tree (Figure 14). For the second, third and fourth trees, the isolate *R. cecicola* was included as an outlier. *R. cecicola* was chosen because it is the closest relative to *Pectinatus* from the superkingdom analysis.

3.7 Genus- and Species-Specific PCR diagnostics

Three sets of PCR primers were designed and tested to specifically detect the genus *Pectinatus*, the species *P. cerevisiiphilus* and the species *P. frisingensis*, respectively. The PCR primer set Pect550 amplifies a 550 bp region of the flagellin gene from both Pc and Pf. The primer sets Pf253 and Pc185 amplify a 253 bp region

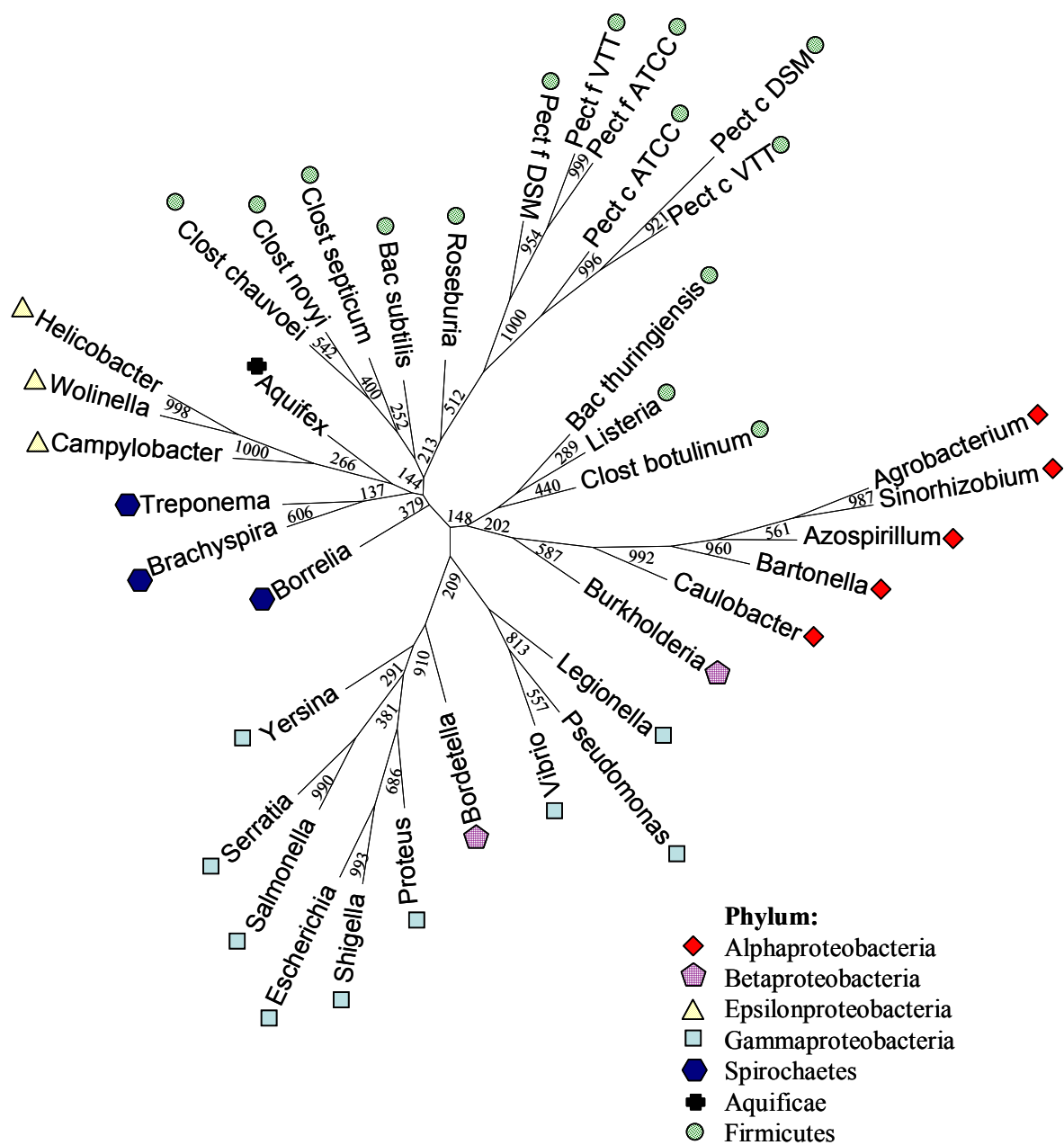


Figure 12 – Phylogenetic tree of flagellin protein sequences across the bacterial superkingdom. Key indicates the current phylum of each species. Bootstrap values are out of 1000 replicates.

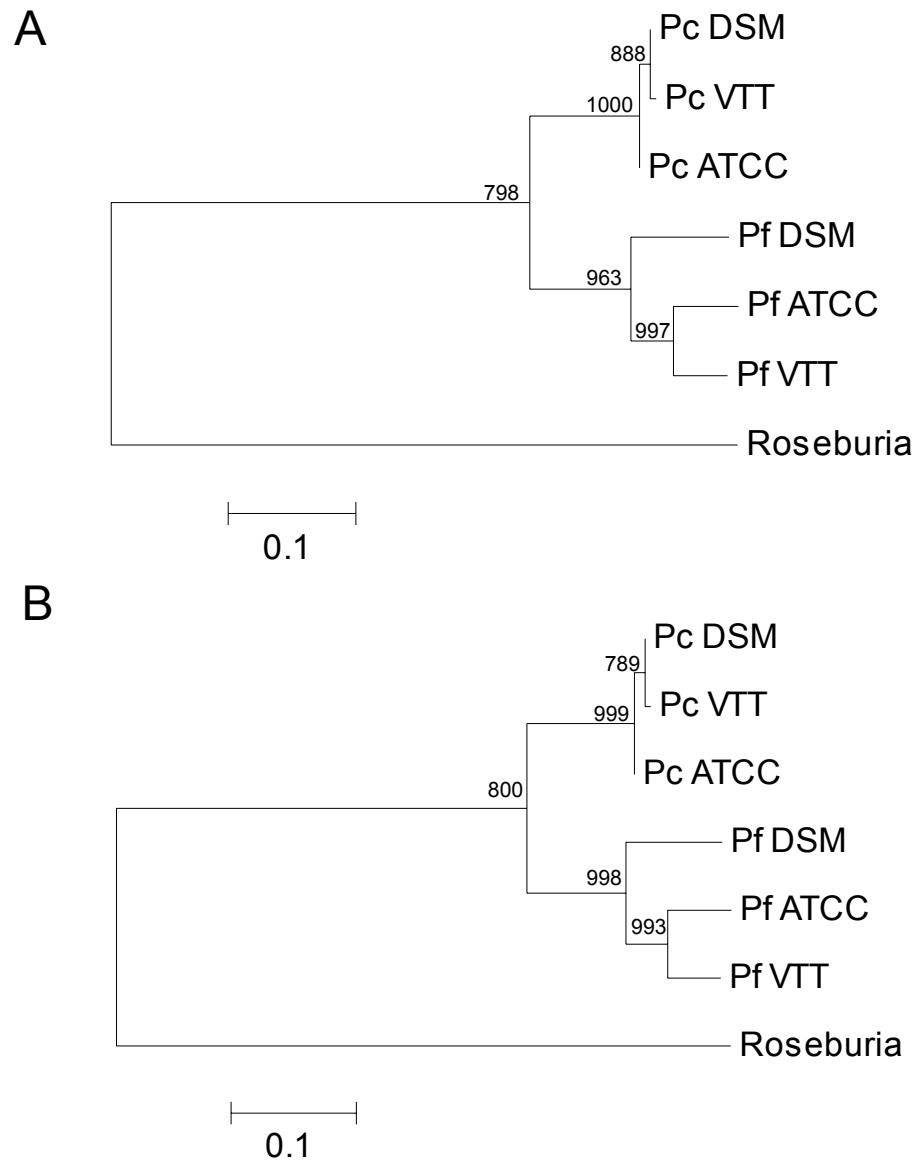


Figure 13 – Phylogenetic trees of *Pectinatus* flagellin at (A) DNA level and (B) protein level. Scale bar indicates number of substitutions per site (1 residue difference per 10 residues). Bootstrap values are out of 1000 replicates.

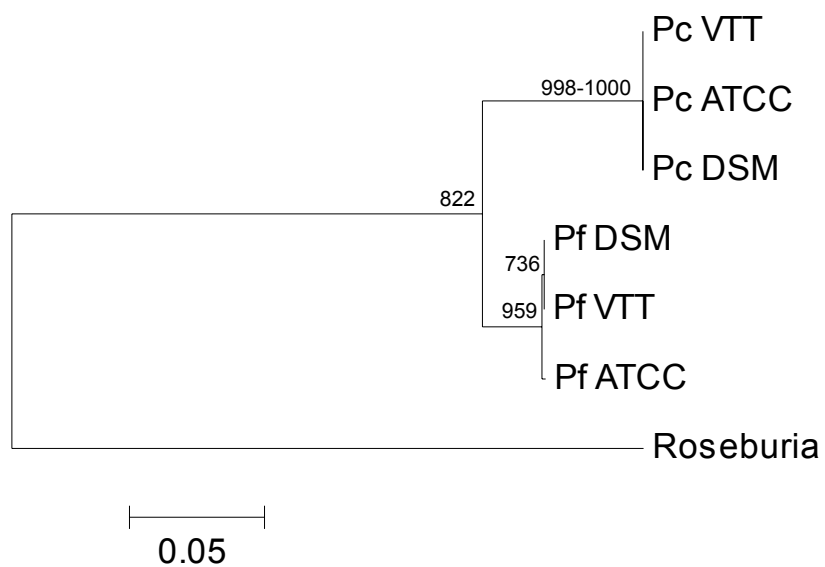


Figure 14 – Phylogenetic tree of *Pectinatus* 16S rRNA sequences. Analysis was based on first three 16S variable regions of the gene. Scale bar indicates number of substitutions per site (1 nucleotide difference per 20 bp). Bootstrap values are out of 1000 replicates.

from Pf flagellin or a 185 bp region from Pc flagellin, respectively. The PCR primer binding sites of all three primer sets are diagramed in Figure 15. All three sets of PCR primers were tested against representative beer spoilage bacteria from all five brewery spoilage classes. Representing common wort spoilers are *Citrobacter freundii* (Cf), *Enterobacter agglomerans* (Ea), *Klebsiella pneumoniae* (Kp), *Obesumbacterium proteus* (Op), *Proteus mirabilis* (Pm) and *Serratia marcescens* (Sm). From the acetic acid bacteria, two *Acetobacter* species (Ace7 and Ace8) and *Gluconobacter oxydans* subsp. *oxydans* (Go) are included. Lactic acid bacteria are represented by four species of *Lactobacillus* (*L. brevis* (L02), *L. delbrueckii* (L14), *L. paracasei* (L20) and *L. plantarum* (L21)) and four species of *Pediococcus* (*P. clausenii* (P06), *P. damnosus* (P12), *P. pentosaceus* (P24) and *P. acidilactici* (P25)). Two *Zymomonas* isolates, *Zymomonas anaerobia* (Za) and *Zymomonas anaerobia* subsp. *anaerobia* (Za [sub a]), are included. Finally, representing the Gram-negative, anaerobic bacteria, aside from the six *Pectinatus* isolates, is *Megasphaera cerevisiae* (Mc). PCR detection results are summarized in Table 8. The primer set Pect550 readily detects five of the six *Pectinatus* isolates (excluding Pf ATCC), weakly detects Ea, Pm and Sm and very weakly detects Cf, Go, Pf ATCC, Za and Za [sub a]. Pf253 only detects Pf DSM, with very weak cross-reactivity with Cf. Pc185 detects all three Pc isolates strongly, Ea, Pf ATCC, Pf DSM and Pm weakly and Cf, Go, Sm, Za and Za [sub a] very weakly. All the other beer spoilage isolates have a negative PCR reaction with the PCR primer sets.

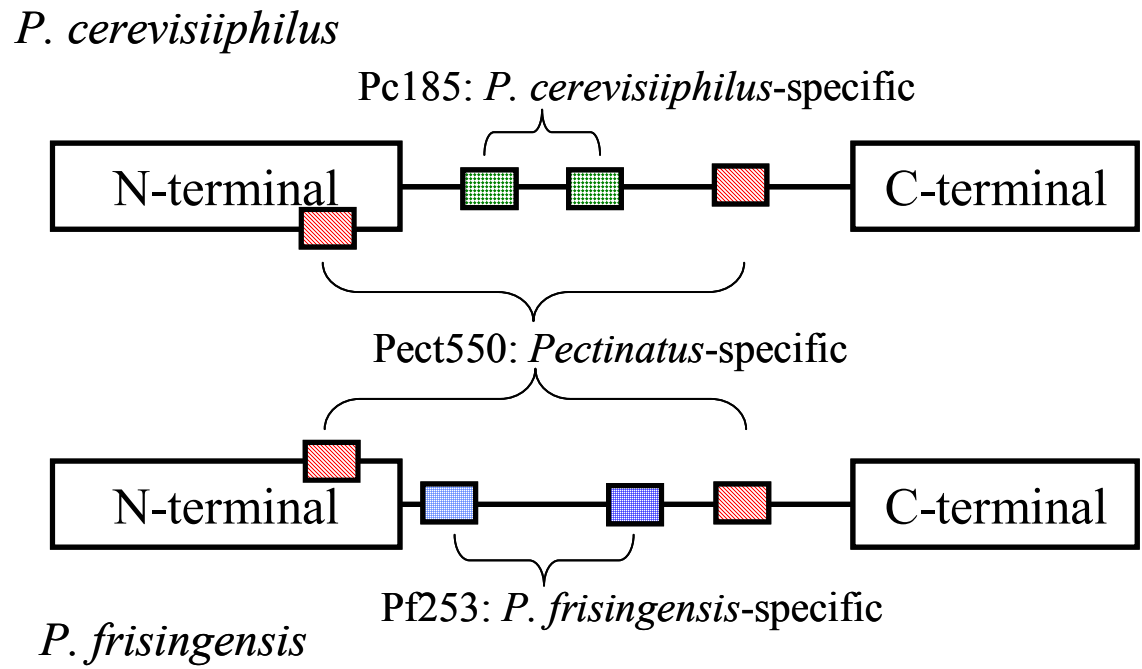


Figure 15 – Binding locations for PCR primers Pect550, Pf253 and Pc185.

Table 8 – Summary of genus- and species-specific PCR results.

Species	Lab Designation	Pect550 ^a	Pf253 ^b	Pc185 ^b
<i>Acetobacter</i> sp.	Ace7	- ^c	-	-
<i>Acetobacter</i> sp.	Ace8	-	-	-
<i>Citrobacter freundii</i>	Cf	-/+	-/+	-/+
<i>Enterobacter agglomerans</i>	Ea	+/-	-	+/-
<i>Gluconobacter oxydans</i> subsp. <i>oxydans</i>	Go	-/+	-	-/+
<i>Klebsiella pneumoniae</i>	Kp	-	-	-
<i>Lactobacillus brevis</i>	L02	-	-	-
<i>Lactobacillus delbrueckii</i>	L14	-	-	-
<i>Lactobacillus paracasei</i>	L20	-	-	-
<i>Lactobacillus plantarum</i>	L21	-	-	-
<i>Megasphaera cerevisiae</i>	Mc	-	-	-
<i>Obesumbacterium proteus</i>	Op	-	-	-
<i>Pectinatus cerevisiiphilus</i>	Pc ATCC	+	-	+
<i>Pectinatus cerevisiiphilus</i>	Pc DSM	+	-	+
<i>Pectinatus cerevisiiphilus</i>	Pc VTT	+	-	+
<i>Pectinatus frisingensis</i>	Pf ATCC	-/+	-	+/-
<i>Pectinatus frisingensis</i>	Pf DSM	+	+	+/-
<i>Pectinatus frisingensis</i>	Pf VTT	+	-	-
<i>Pediococcus acidilactici</i>	P25	-	-	-
<i>Pediococcus clausenii</i>	P06	-	-	-
<i>Pediococcus damnosus</i>	P12	-	-	-
<i>Pediococcus pentosaceus</i>	P24	-	-	-
<i>Proteus mirabilis</i>	Pm	+/-	-	+/-
<i>Serratia marcescens</i>	Sm	+/-	-	-/+
<i>Zymomonas anaerobia</i>	Za	-/+	-	-/+
<i>Zymomonas anaerobia</i> subsp. <i>anaerobia</i>	Za [sub a]	-/+	-	-/+

^a Pect550 primer set was used with an annealing temperature of 59°C.

^b Pf253 and Pc185 primer sets were used with annealing temperatures of 61°C.

^c + ; positive, +/- ; weak positive, -/+ ; very weak positive, - ; negative.

4.0 DISCUSSION

4.1 *Pectinatus* flagellin

4.1.1 Sequencing the flagellin

Identifying and sequencing the flagellin genes from *Pectinatus* was done in a number of steps. The first step was to identify at least two amino acid sequences from each flagellin protein. These peptide sequences were needed to design degenerate PCR primer sets to amplify a section of each flagellin gene. Micro-protein sequencing was successful in generating two amino acid sequences from Pc ATCC (the original N-terminal and an internal peptide fragment), but only one amino acid sequence from Pf DSM (the original N-terminal) (Figure 4). Sequencing of two Pf DSM flagellin peptide fragments was attempted, but both resulted in mixed sequences (more than one aa at each position). This was likely due to the formation of two or more peptides that had approximately the same molecular weight, causing them to be isolated together. Without a second peptide sequence for Pf DSM flagellin, PCR primer design and sequencing of this isolate was delayed until the flagellin gene from Pc ATCC was sequenced. It was believed that flagellin sequence similarities between the two *Pectinatus* species and knowledge of the conserved domains found in all bacterial flagellin would be enough to overcome the need for a second Pf DSM specific amino acid sequence. This proved correct.

The second step in sequencing the flagellin genes was to reverse translate the amino acid sequences from Pc ATCC into degenerate PCR primers. The forward PCR primer, Flgn2F, was designed from the aa sequence of the original protein N-terminal, while the reverse PCR primer, Flgn3R, was designed from the peptide fragment aa sequence. Because it was unknown exactly where the Pc ATCC peptide fragment fit into the intact flagellin protein, a PCR product size using the primers Flgn2F and Flgn3R could not be predicted *a priori*. The PCR primer combination generated a 419 bp product, which corresponded to most of Pc ATCC's N-terminal domain (Figure 5). This was the specific sequence needed to begin Bubble-PCR sequencing of the rest of the gene.

The third step in sequencing the flagellin genes was to use a combination of PCR and Bubble-PCR to finish sequencing Pc ATCC and sequence Pf DSM. Bubble-PCR was used to “genome walk” upstream and downstream of the N-terminal region of the Pc ATCC flagellin gene. Once the flagellin gene sequence was completed for Pc ATCC, various Pc ATCC PCR primer combinations were tested with Pf DSM genomic DNA. The expectation was that at least one primer set would be at positions that shared sequence identity between the two species. The PCR primer set Flgn8F and Cterm1 met these conditions. Flgn8F is a Pc ATCC-specific primer located at the edge of the C-terminal flagellin domain and Cterm1 is a degenerate primer designed from all bacterial flagellin genes. Together, this primer set amplified the C-terminal domain of the Pf DSM flagellin gene, creating a new starting point for Bubble-PCR. A combination of PCR and Bubble-PCR was then used to sequence the flagellin gene from Pf DSM.

Bubble-PCR has advantages and disadvantages as a gene identification and sequencing method. Once a specific region of the target gene has been identified,

Bubble-PCR is a straightforward and rapid method of sequencing. This is because Bubble-PCR is a variation of PCR, meaning the reaction time is short, equipment needed is minimal and the procedure is not technically difficult. The limiting factor of this method is the speed with which PCR products can be sequenced and new PCR primers generated. The drawback of Bubble-PCR is that a specific sequence is needed to create the first gene-specific PCR primer. Bubble-PCR was attempted with the degenerate PCR primers FlgnTerminal1 and Flgn10F as the gene-specific primer without success. One reason for this result might be in the annealing temperature used for the reaction. With a degenerate primer, an annealing temperature of 40⁰C is used (to allow for bp mismatch tolerance), while for a specific primer, an annealing temperature of 60⁰C is used (to ensure primer binding specificity). It is possible that the primer BubblePrimer, which is 30 bp, requires the higher temperature to maintain proper binding to its target sequence only. It is also worth noting that in this study, the primer Flgn13R did not work with the Bubble-PCR system. This may indicate that there are limitations within Bubble-PCR for specific PCR primers as well. There is no apparent difference between the primer Flgn13R and other Flgn primers. As such, it is unknown whether the reason for the failure of Flgn13R primer is primer- or method-based.

With Pc ATCC and Pf DSM flagellin genes sequenced, the final step was to sequence two additional isolates from each species, namely Pc DSM, Pc VTT, Pf ATCC and Pf VTT. By comparing the two completed *Pectinatus* flagellin sequences, it was noted that the first and last 22 bp of the open reading frames are identical. Concurrently, a PCR primer set (PectCloneF and PectCloneR) was designed to these regions for the full-length amplification of Pc ATCC and Pf DSM flagellin for cloning experiments. This PCR primer set was tried for the four unsequenced flagellin genes and was found to

yield the expected PCR products. This allowed the flagellin genes from the four isolates to be sequenced without the need for Bubble-PCR. However, the drawback to this approach was that the PectClone primer set bound within the open reading frames of the flagellin genes. When the PectClone primers were used to sequence the flagellin PCR products, the DNA sequence under and immediately downstream of the primers was not able to be determined. This is why the flagellin gene sequences for Pc DSM, Pc VTT, Pf ATCC and Pf VTT are incomplete at their 5' and 3' ends.

4.1.2 Characteristics of the flagellin open reading frame

The complete flagellin sequence was only determined for Pc ATCC and Pf DSM, leaving Pc DSM, Pc VTT, Pf ATCC and Pf VTT with sequence gaps at the beginning and end of the coding sequence. Although Pc DSM and Pc VTT are incomplete at the N- and C-terminals, all of the sequences to date for these isolates are identical to Pc ATCC. For the sake of this discussion, all *P. cerevisiiphilus* flagellins are considered to have an identical sequence over their entire length. Conversely, *P. frisingensis* flagellin genes show variation throughout their sequences, so any discussion of Pf ATCC and Pf VTT sequence refers only to the actually sequenced regions. Discussions referring to the overall *Pectinatus* flagellin genes are made in reference to Pc ATCC and Pf DSM.

The open reading frames of the flagellin genes from *P. cerevisiiphilus* and *P. frisingensis* code for proteins 446 aa and 448 aa, respectively. These sizes falls within the 276 aa to 575 aa range of known flagellin proteins (Wilson and Beveridge, 1993). The two aa difference between Pc and Pf is found at the same position within the variable region for all sequences (Figures 8 and 9). It is unknown whether this gap represents an insertion in the Pf flagellin sequence or a deletion in the Pc flagellin

sequence. Either way, this distinct characteristic is one way to differentiate between the two species of *Pectinatus*.

Using the National Center for Biotechnology Information Conserved Domain Database, all six *Pectinatus* flagellin proteins are found to contain flagellin domains listed in both the Pfam (protein families) and COG (clusters of orthologous groups of proteins) databases (Marchler-Bauer *et al.*, 2003). Pfam is a large collection of multiple sequence alignments and hidden Markov models covering many common protein families. COGs, on the other hand, were delineated by comparing protein sequences encoded in 43 complete genomes, representing 30 major phylogenetic lineages. Each COG consists of individual proteins or groups of paralogs from at least three lineages and thus corresponds to an ancient conserved domain (Tatusov *et al.*, 2001). From the Pfam database, the N-terminal flagellin domain (pfam00669.8) and the C-terminal flagellin domain (pfam00700.8) are recognized. From the COG database, the flagellin and related hook-associated proteins group, FlgL (COG1344.1) is found. These database matches reinforce the fact that *Pectinatus* flagellin proteins share a common ancestry with all bacterial flagellin. This relationship also justifies phylogenetic analysis as a method for comparison, since phylogeny reflects the evolutionary relatedness of sequences (Brinkman and Leipe, 2001).

The amino acid composition of *Pectinatus* flagellin proteins is also consistent with established bacterial flagellin proteins. Pc flagellin proteins contain a high percentage of alanine residues (10.3%) and low percentages of histidine (0.4%), proline (0.2%) and tryptophan (0.2%). There are no cysteine residues in the Pc flagellin sequences. Similarly, Pf DSM flagellin proteins contain 11.6% alanine and 0.2% tryptophan and proline. No histidine or cysteine residues are found in the sequence.

Finally, Pf ATCC and Pf VTT have 10% and 9.8% alanine, respectively, and are both comprised of 0.2% histidine, proline and tryptophan. Once again, neither flagellin sequence contains any cysteine residues. This characteristic amino acid composition fits well with other characterized flagellins (Wilson and Beveridge, 1993).

The cloning and expression of Pc ATCC and Pf DSM flagellin genes were done as GST-fusion proteins in *E. coli*. The predicted size of both flagellin proteins from their translated sequences is about 47 kDa. By adding a GST tag (26 kDa) to the beginning of the flagellin protein, the theoretical size of the GST-flagellin fusion protein is approximately 73 kDa. Both anti-*Pectinatus* flagellin antibodies and anti-GST antibodies detected the fusion proteins of both species at approximately 68 kDa (Figure 7). This is within a reasonable margin to the theoretical size. Obtaining the theoretical size means that the fusion proteins are not post-translationally modified by *E. coli*. This is not surprising, since *E. coli* does not post-translationally modify its own flagellin proteins (Kuwayama, 1988). The ladder-like banding pattern of the fusion proteins indicates that both proteins are degraded by the host *E. coli* cells (Figure 7, lanes 2 and 3). This is an expected phenomenon when a foreign protein is introduced into a cell (Makrides, 1996). Finally, an additional finding is that anti-*Pectinatus* flagellin antibodies do not cross-react with *E. coli* flagellin proteins, evident by the absence of bands in the corresponding pGEX control lane (Figure 7, blot A, lane 4).

4.1.3 Regulatory elements

DNA sequencing upstream and downstream of Pc ATCC and Pf DSM coding sequences reveal all the expected regulatory elements of a bacterial flagellin gene. Upstream of the start codon are the promoter and ribosomal binding sites. The promoter

used by bacterial flagellin is the sigma 28 promoter. First characterized in *B. subtilis*, this promoter has been found with many flagellin genes (Figure 10). The putative sigma 28 promoter found in *Pectinatus* species shares the same general sequence and arrangement as other characterized flagellin, with one exception. The Pf DSM sequence has 16 bp between the -10 and -35 motifs of the promoter, instead of the usual 15 bp (Figure 10). The only other known flagellin gene to have this extra bp in the promoter is *Clostridium difficile* (Tasteyre *et al.*, 2000). It is currently unknown what affect, if any, this increased space has on the ability of the sigma 28 subunit to bind the promoter sequence. Downstream of the promoter, and approximately 10 bp upstream of the start codon, is the ribosomal binding site. Also known as the Shine-Dalgarno site, this region of DNA has the sequence AGGAGGU, which is complementary to part of the 16S ribosomal subunit. This site allows the ribosome to position itself at the start of the open reading frame and begin translation (Snyder and Champness, 1997). Most flagellin genes have short, four base pair Shine-Dalgarno sites (Kuwajima *et al.*, 1986; Guerry *et al.*, 1990; Wallich *et al.*, 1990; McGee *et al.*, 1996; Spangenberg *et al.*, 1996). By comparison, the ribosomal binding site AGGAGG in Pc ATCC represents almost the entire Shine-Dalgarno sequence. Pf DSM contains one less conserved bp than Pc ATCC, with the sequence GGAGG. Both ribosomal binding sites allow for the strong and specific association of the ribosome with the flagellin mRNA transcripts.

Downstream of the open reading frame, a termination loop is expected. The termination loop is a palindromic sequence that, when single-stranded, will fold together to form a stem-loop structure. A termination loop is always required for rho-independent termination and is commonly found in rho-dependent termination. Only Pf DSM was sequenced downstream far enough to detect a palindromic sequence. Within

the Pf DSM sequence, a perfect stem of 17 bp is found with an intervening loop of 2 bp (Figure 8, 1625-1660 bp). The stem sequence contains 12 A+T pairs and 5 G +C pairs. Since this stem-loop structure is not G+C rich and there is no tailing A stretch, the transcription termination for this gene is most likely in a rho-dependent fashion.

4.1.4 Glycosylation

As stated previously, the theoretical size of all *Pectinatus* flagellin proteins is about 47 kDa. As determined by SDS-PAGE, the apparent size of extracted flagellin proteins from intact *Pectinatus* cells is from 50 to 60 kDa. The size difference between the theoretical and apparent flagellin proteins from *Pectinatus* species is due to the post-translational modification, glycosylation. General detection of carbohydrates was done via a number of carbohydrate-specific staining procedures. Attempts to determine the identity of the terminal sugar group were made with specific stains and endoglycosidases. Finally, the type of carbohydrate linkage to the protein backbone was investigated with an exoglycosidase.

The general carbohydrate detection staining results are summarized in Table 7 and Figure 11. Ova and HRP were used as positive controls while BSA was used as a negative control (Montgomery, 1970). Both PAS staining and thymol-sulfuric acid staining are able to detect carbohydrate on five of the six *Pectinatus* flagellin (excluding Pf ATCC). The bands from PAS staining were bright and of equal intensity for Pc ATCC, Pc DSM, Pc VTT, Pf DSM and Pf VTT, indicating that approximately the same number of 1, 2-diol groups were visualized from each glycoprotein. Conversely, the bands generated for Pf DSM and Pf VTT from thymol-sulfuric acid staining were approximately 1/10th the intensity of other bands. This suggests that there are fewer C-2

hydroxyl containing carbohydrates on the flagellin proteins from the two Pf isolates as compared to the Pc isolates. Exactly why Pf ATCC did not stain with either method is unknown. Since it is unlikely that another type of post-translational modification is present only in Pf ATCC, the simplest answer is that Pf ATCC contains carbohydrate moieties that do not contain either 1, 2-diol or C-2 hydroxyl groups. There are examples in the literature where glycosylation was suspected for flagellin proteins, but was undetectable by current methods (Brimer and Montie, 1998). Finally, the stain alcian blue, which histologists use to detect acidic glycoproteins in human tissue, was also tried. Unfortunately, the stain was incompatible with SDS-PAGE protein separation and all proteins tested were visualized.

Once the presence of carbohydrate on the flagellin proteins was confirmed, a number of methods were employed to determine the identity of the terminal sugar residue. The terminal sugar tested for most extensively was the family of sialic acids. Sialic acid is a generic term for both N- and O-acetylated neuraminic acids. Sialic acids commonly occur as the terminal residue of carbohydrate chains in many glycoproteins and glycolipids. To determine the presence of sialic acids on *Pectinatus* flagellin, the type strain of each *Pectinatus* species was subjected to two procedures; PAS staining specific for sialic acid and neuraminidase cleavage. The glycoprotein Fet, which is known to contain approximately 8% sialic acid, was used as a positive control (Graham, 1972). PAS staining was made specific for sialic acids by reducing the concentration of periodic acid used (Hounsell *et al.*, 1996). SDS-PAGE gels of separated proteins stained by this method did not reveal the presence of sialic acids on *Pectinatus* flagellin. The absence of sialic acid on *Pectinatus* flagellin was confirmed by neuraminidase treatment. Neuraminidase, a type of endoglycosidase, hydrolyses the terminal α -2 \rightarrow 3, 2 \rightarrow 6 and

2 → 8-linkages which bind N-acetylneuraminic acid to N-acetylhexosamines and N- or O-acetylated neuraminyl residues in oligosaccharides, glycolipids and glycoproteins (Uchida *et al.*, 1979). If sialic acids were present on the flagellin proteins, treating them with neuraminidase before SDS-PAGE would cause a decrease in their apparent molecular weight. This is because the sialic acid residues would be removed prior to SDS-PAGE separation and the overall size of the glycoprotein would be smaller. Treatment of Pc ATCC and Pf ATCC with neuraminidase resulted in no such shift. Therefore, based on both PAS staining and neuraminidase cleavage, it is concluded that *Pectinatus* flagellin do not contain sialic acids. In this study, the endoglycosidases α- and β-galactosidase (removes both forms of galactose), α-fucosidase (removes fucose) and N-acetyl-β-D-glucoaminidase (removes GlcNAc) were also tested. None of these enzymes had any effect on Pc ATCC or Pf ATCC. However, known glycoproteins containing these terminal sugars were not obtained and tested with these enzymes. As such, without positive controls to confirm the activities of the endoglycosidase, it is premature to conclude that *Pectinatus* flagellin proteins do not contain any of these terminal sugar residues.

The final property of *Pectinatus* flagellin glycosylation that was investigated was the type of carbohydrate-protein linkage. As stated previously, carbohydrate can be linked to proteins through one of two linkages: N-linked or O-linked. Unfortunately, only exoglycosidases that remove N-linked carbohydrate are available. To remove N-linked carbohydrates, the exoglycosidase PNGase F is used. Pc ATCC and Pf ATCC flagellin proteins were digested with PNGase F and separated on SDS-PAGE to visualize any changes in their apparent molecular weight. Both Ova and Fet are known to be N-linked and were used as positive controls (Montgomery, 1970; Graham, 1972).

Both *Pectinatus* type stains are resistant to PNGase F cleavage, indicating that they are not glycosylated through an N-linkage. One noteworthy exception to the ability of PNGase F to cleave N-linkages is for a carbohydrate containing a core α 1-3 fucose. This type of carbohydrate is rare and is not usually a major consideration. However, Senchenkova *et al.* (1995) determined that both *P. cerevisiiphilus* and *P. frisingensis* contain fucose sugars in their LPS layers. Whether fucose would also be utilized for glycosylation is unknown, but the presence of this rare sugar in *Pectinatus* should not be forgotten. As such, it is probable, but not absolute, that *Pectinatus* flagellin proteins contain O-linked carbohydrates.

Within a discussion of carbohydrate linkage types, a brief mention of related predictive software is warranted. The large datasets accumulated from genomics and proteomics projects have resulted in a wide variety of predictive (bioinformatics) software being developed. Without the availability of an exoglycosidase to investigate O-linked glycoproteins, the glycosylation site predictive program NetOGlyc has emerged (Hansen *et al.*, 1998). NetOGlyc is a collection of artificial neural networks that has been trained to recognize eukaryotic mucin type O-glycosylation sites based on sequence context and surface accessibility derived from amino acid sequence. When bacterial flagellins were submitted to NetOGlyc 2.0 (available at <http://www.cbs.dtu.dk/services/NetOGlyc/>), the program had low predictability for experimentally determined flagellin O-glycosylation sites. This made NetOGlyc unsuitable for generating accurate predictions pertaining to bacterial flagellin. Poor performance from NetOGlyc was not surprising considering that the neural network was trained to recognize eukaryotic, not prokaryotic, glycoproteins. The major obstacle in developing a prokaryotic equivalent to NetOGlyc is the limited number of

experimentally characterized bacterial glycoproteins currently available compared to the relatively large number of glycoproteins needed to adequately train such a network. As basic research continues into bacterial O-linked glycoproteins, this obstacle will hopefully be removed and a prokaryotic equivalent to NetOGlyc will be available as an experimental alternative.

4.2 Phylogeny

To gain a better understanding of the genus *Pectinatus*, the flagellin genes were subjected to phylogenetic analyses. Phylogenetic analysis is the study of evolutionary relationships (Brinkman and Leipe, 2001). The conserved nature of the flagellin genes, both at sequence and functional levels, lend themselves to this type of analysis. Although the 16S rRNA gene is the most common sequence used for phylogenetic analyses, flagellin genes have been studied successfully (Fukunaga and Koreki, 1996; Winstanley and Morgan, 1997). It is also reassuring that phylogenies inferred from molecules other than rRNA genes have been found to be in good agreement with 16S rRNA gene analyses (Olsen and Woese, 1993; Dobson, 2001).

There are many methods for conducting phylogenetic analyses. Good reviews of the advantages and disadvantages of current methods can be found in Brinkman and Leipe (2001) and Dobson (2001). For the scope of this discussion, descriptions and justifications will be limited to the methodologies used. In this study, sequences were aligned with ClustalX (Thompson *et al.*, 1994; 1997) and manually inspected for conformity to known domains with GeneDoc (Nicholas *et al.*, 1997). Based on the current multiple alignment techniques, this approach is regarded as among the best currently available (Brinkman and Leipe, 2001). From the multiple alignments, some

measure of sequence relatedness must be determined. This is done by creating a distance matrix from calculations of how similar any DNA or protein sequence is to any other. To calculate the distances between DNA sequences, the Kimura 2-parameter matrix was used, while for protein sequences, the Dayhoff PAM matrix was chosen. The Kimura 2-parameter matrix sets independent rates for transitions and transversions and, as such, is a realistic approximation of actual DNA mutation events (Kimura, 1980). The Dayhoff PAM matrix is based on an empirical model that calculates probabilities of one amino acid being changed to another (Dayhoff *et al.*, 1978). For protein sequences with reasonable similarity, like flagellin sequences, the PAM matrix has been shown to be well suited (Brinkman and Leipe, 2001).

After creating a distance matrix, the next step is to build a phylogenetic tree. From the many tree-building methods available, this study used distance tree building using neighbor joining. This type of tree construction involves using the values from the distance matrix to group the two most similar sequences together first. These two sequences are then combined and considered as one. Another round of comparison is done and the two most similar sequences are again grouped. This procedure is repeated until all the sequences in the analysis are grouped together. Neighbour joining was used to construct the phylogenetic trees because it is the fastest tree building method and generally yields a tree close to the minimum evolution tree (a more computationally intensive method that currently appears to be the best procedure) (Brinkman and Leipe, 2001). Neighbour joining is considered less appealing than other methods because it only generates one tree. This limitation was overcome by conducting bootstrap analyses of all trees. Bootstrapping is a statistical re-sampling technique that involves randomly changing positions in the initial multiple sequence alignment. This creates multiple,

slightly different data sets. A distance matrix and phylogenetic tree are constructed for each data set and then all the trees are compiled into one consensus tree. This consensus tree contains bootstrap values that indicate how many times each consensus branch was found in the individual data set trees. The more often a branch is conserved, the higher the probability that the branch is correct. In this study, bootstrapping was done with 1,000 replicates. Using the methods described, phylogenetic trees were constructed involving the *Pectinatus* flagellin.

4.2.1 Bacteria (superkingdom) level

Four phylogenetic analyses were done to obtain taxonomic data about the genus *Pectinatus*. The first analysis compared the six *Pectinatus* flagellin proteins to other flagellin proteins from 30 species throughout the bacterial superkingdom (Figure 12). This tree is a consensus tree, and as such, the branch lengths depicted are equal for all species. The length has no evolutionary meaning. The intent of this analysis is to represent the pattern of divergence, not the distance of divergence. The overall tree topography is in agreement with the analysis of flagellin genes done by Winstanley and Morgan (1997). The *Pectinatus* isolates clustered into their two species 100% of the time (evident by the bootstrap value of 1,000). The *Pectinatus* branch is found clustered with the low G+C, Gram-positive phylum *Firmicutes*. This data supports the recent move of *Pectinatus* to this phylum. The closest neighbour to *Pectinatus*, based on the branching pattern of the phylogenetic tree, is *R. cecicola*. Like *Pectinatus*, *R. cecicola* is also a Gram-negative, anaerobic bacterium that groups with Gram-positive organisms based on 16S rRNA sequence (Martin and Savage, 1988). In fact, besides *Roseburia* and *Pectinatus*, a number of species including *Megasphaera*, *Selenomonas* and

Sporomusa are Gram-negative by morphology, but contain Gram-positive 16S rRNA signatures (Stackebrandt *et al.*, 1985). These genera represent an intermediate between the two arbitrary classes of bacteria. It is significant to note that although the flagellin gene represents a structural component of the bacteria, its sequence still agrees with the phylogeny of the 16S rRNA gene. This makes the flagellin gene a useful taxonomic marker outside of the 16S rRNA region. General inspection of the rest of the tree confirms that branching patterns generally agree with the current taxonomy of eubacteria.

4.2.2 *Pectinatus* (genus) level

The other three phylogenetic analyses conducted were comparisons within the *Pectinatus* genus using the flagellin DNA and protein sequences (Figure 13A and 13B, respectively) and the 16S rRNA gene sequence (Figure 14). For all three trees, *R. cecicola* was included as an outlier and both branch lengths and bootstrap values were calculated. The flagellin trees were compared to the 16S rRNA tree to determine the compatibility of the two genes for phylogenetic assignment. Because DNA mutations can go unnoticed at the protein level, especially at the third bp position in a codon, both the DNA and protein sequences from *Pectinatus* flagellin were examined. The tree topography is identical for the DNA and protein levels of the flagellin. Both trees illustrate that the Pc isolates studied have diverged very little, while the Pf isolates have more variation in their sequence. Conversely, when compared to the outlier *Roseburia*, *Pectinatus* isolates still represent a closely related genus. The flagellin genes also embody much more sequence variation than the 16S rRNA gene region, as evident by the longer branch lengths found in the flagellin trees. Indeed, the 16S rRNA gene is

identical for all three Pc isolates and had only one nucleotide difference for Pf ATCC among the Pf isolates. This reinforces the fact that the 16S rRNA genes are highly conserved and emphasizes the need to examine genes outside the rRNA region to understand the taxonomic relationships of closely related organisms.

4.3 Detection of *Pectinatus* via the flagellin gene

One of the primary goals of this study is to determine if the flagellin gene from *Pectinatus* would make a suitable target for a PCR-based rapid detection protocol. To this end, three PCR primer sets were generated. The first set, Pect550, is designed to specifically detect isolates belonging to the genus *Pectinatus*. The second set, Pf253, is designed to detect only Pf isolates and the third set, Pc185, is designed to detect only Pc isolates. At the time of PCR primer design, there was a major limiting factor. Only the flagellin sequences from Pc ATCC and Pf DSM had been determined. As such, Pect550 is modeled after only two sequences, while Pf253 and Pc185 are based on only one sequence each. It was known that genus- and species-specific PCR primers were unlikely to be completely effective based on such a small sample size. The primer sets were not redesigned when Pc DSM, Pc VTT, Pf ATCC and Pf VTT flagellin sequences were completed due to the fact that several additional *Pectinatus* isolates are expected shortly. A complete revision of possible genus- and species-specific primers will await this much larger data set. To redo PCR primers to account for six *Pectinatus* isolates and then redo the analysis again for a much larger number of isolates was deemed an unnecessary duplication. Therefore, the results presented here represent a preliminary attempt at a *Pectinatus*-specific detection system. The PCR detection results, listed in Table 8, reveal the properties of each of the three primer sets that were tested. The

primer set Pect550 readily detects five of the six *Pectinatus* isolates (excluding Pf ATCC), with cross-reactivity to a number of wort spoilers, Go and *Zymomonas* isolates. The primer set Pf253 is isolate specific for Pf DSM (the strain it was designed to), having only a minor reaction with Cf. The primer set Pc185 detects the three very closely related Pc isolates, but also detects two of the three Pf isolates reasonably well, along with cross-reactivity to other beer spoilage organisms. Based on these results, all three primer sets detect the isolate(s) they were designed to, but did not expand well to their respective target groups exclusively.

The PCR results show very encouraging signs that a brewery detection system for *Pectinatus* can be based on the flagellin gene. None of the three primer sets reacted with any of the lactic acid bacteria (*Lactobacillus* and *Pediococcus*) or *Megasphaera*. This is expected since none of these organisms express flagella. This illustrates the advantage of targeting a gene that is not universally found in bacteria, unlike the 16S rRNA gene or 16S-23S interspacer region. By selecting the flagellin gene as the detection target, the possibility of cross-reaction with non-flagellated bacteria is eliminated completely. Another consideration to remember when evaluating the PCR detection results presented in Table 8 is that all beer spoilage organisms will not be present in the same samples. In the brewery, *Pectinatus* contamination is not found in wort nor are wort spoilage organisms found at the packaging stage. Different environmental and nutritional requirements limit beer spoilage organisms to specific locations in the brewery. So, although academically it is desirable to create a *Pectinatus* detection system that does not cross-react with any other brewing spoilage organism, it is actually not necessary to achieve such a goal. In practical terms, a *Pectinatus* detection system only needs to be able to detect *Pectinatus* isolates to the exclusion of any other

microorganisms that might be present in the same sample. To date, this has been achieved.

Two things must be done before a detection protocol based on *Pectinatus* flagellin genes can be implemented. The first is to redesign genus- and species-specific PCR primers with an enlarged sample set. This will be done by the sequencing and comparison of flagellin genes from more *Pectinatus* isolates. An enlarged sample size may reveal that one PCR primer set may not be able to detect all the isolates of a designated group at the exclusion of other groups. One way to overcome a limitation like this is to design sets of isolate-specific primers (like the primer set Pf253 turned out to be) and pool select groups of them to achieve the desired specificity. The second thing to be done before implementing new primer sets is to expand the test organism set to include the anaerobic beer spoilage genera *Selenomonas* and *Zymophilus*. These organisms are taxonomically close to *Pectinatus* and have caused cross-reactivity problems in the past (Satokari *et al.*, 1998). For a flagellin-based PCR detection system to out-perform current *Pectinatus* PCR detection systems, isolates from *Selenomonas* and *Zymophilus* must be excluded from detection. This will probably mean that the flagellin genes from these genera will also have to be sequenced for comparison. When these two tasks are successfully completed, the resulting PCR protocol will allow for the rapid and specific detection of *Pectinatus* isolates in the brewery.

4.4 Conclusions

Twenty-five years ago, *Pectinatus* isolates became recognized as beer spoilage bacteria. Today, they continue to be a problem for packaged, unpasteurized beer. Research into *Pectinatus* has focused primarily on physical and growth characteristics,

as well as detection strategies for these organisms. Very little work has been done to uncover and understand the genetics of this genus. This study set out to change that by sequencing and characterizing the flagellin genes and proteins of *Pectinatus*. This information was then used to study the evolutionary relationships of *Pectinatus* isolates, both compared outward to other bacterial flagellin and inward to relationships at the genus and species levels. Finally, the new flagellin information was applied to the problem of rapidly detecting *Pectinatus* isolates in the brewery by conducting the preliminary testing for a PCR-based detection protocol.

The sequencing and characterization of the flagellin genes revealed many similarities to known bacterial proteins. All the expected regulatory elements for a bacterial flagellin gene are present in *Pectinatus*, including a sigma 28 promoter, -10 ribosomal bind site and terminating loop. *Pectinatus* flagellin proteins are 446 aa (for Pc) or 448 aa (for Pf) long. This is well within the size range of known bacterial flagellin. *Pectinatus* flagellin proteins also share the same general aa composition and domain structure as other characterized flagellins. Finally, sharing the trait of a few other well characterized flagellins, *Pectinatus* flagellin proteins are glycosylated.

The glycosyl moieties on *Pectinatus* flagellin add 5-10 kDa to the apparent molecular weight of the protein. The carbohydrates are most likely linked to the protein backbone through an O-glycosyl linkage. Although the carbohydrate make-up of the post-translational modification was not determined, the terminal sugar on *Pectinatus* is not from the sialic acid family. As well, three different combinations of carbohydrates are present among the six isolates tested. All three Pc isolates have both 1, 2-diol groups and C-2 hydroxyl groups present as part of their terminal sugars. Pf DSM and Pf VTT also have both groups, but the C-2 hydroxyl groups are present at a much lower

abundance. Pf ATCC did not contain either sugar structure. In fact, Pf ATCC was negative for all carbohydrate detection methodologies. Since it is unlikely that Pf ATCC is modified by a completely different process, the carbohydrate moiety on Pf ATCC is such that it is undetectable by the methods employed in this study. The glycosylation of *Pectinatus* flagellin proteins appears to be an organism-specific event, since the GST-flagellin fusion proteins generated in *E. coli* were not glycosylated.

The phylogenetic analyses conducted with the *Pectinatus* flagellin proteins revealed relationships both outside and inside the genus. The comparison of *Pectinatus* flagellin to other bacterial flagellin agrees with previously described relationships that state that *Pectinatus* belongs with the phylum *Firmicutes* (Gram-positive, low G+C bacteria). This is an interesting finding. The morphology of *Pectinatus* classifies this genus as Gram-negative, while the 16S rRNA signature classifies it as Gram-positive. It is interesting to find that the flagellin gene, which codes for a structural protein, agrees with the organism's classification based on the 16S rRNA gene instead of the classification expected from its structural features. This adds weight to the argument that a consistent bacterial taxonomy can be determined from DNA sequences. Within the *Pectinatus* genus, phylogenetic analysis shows that the Pc isolates studied are very homogeneous, while the Pf isolates contain more heterogeneity. Comparing the phylogenetic trees based on flagellin sequences versus trees based on 16S rRNA sequences, more evolutionary distance was apparent in the flagellin sequences. This likely reflects the ability of protein coding genes to tolerate more sequence variation without altering their resulting protein's function. This finding also illustrates the utility of using protein coding genes to investigate the phylogenies of very closely related organisms.

The sequence information from *Pectinatus* flagellin genes was applied to the problem of rapid detection of this genus in the brewery setting. The preliminary genus-specific PCR primer set, Pect550, and the species-specific PCR primer sets, Pf253 and Pc185, were unable to detect only their designated groups. Pect550 and Pc185 had problems with cross-reactivity to other beer spoilage bacteria. Conversely, Pf253 was too specific and only detected the isolate it was designed to. The results from these PCR primer sets do offer promise for future work. *Pectinatus* isolates gave very bright bands as positive results while cross-reacting bands were always weak to very weak. More work in this area will determine if a rapid detection protocol can be designed from the *Pectinatus* flagellin genes.

4.5 Future considerations

Future considerations for *Pectinatus* research follow the three major veins of study thus far. The first is to continue studying glycosylation of the flagellin genes. There are very few proteins in bacteria that are glycosylated and study of this phenomenon is generally reserved for eukaryotes. An understanding of how bacteria glycosylate their proteins could lead to a better understanding of this type of post-translational modification in general. Specific to *Pectinatus*, discovering the sugars and linkages present in the glycosyl moieties could lead to new detection methods. Lectins, which are proteins that bind specific carbohydrate linkages, could replace antibodies as the specific molecule needed to detect *Pectinatus* organisms in current immunoassays. This could make some techniques, like membrane filter-based fluoroimmunoassays, more practical by removing the costly and time-consuming job of making and characterizing antibodies. Whether through direct binding of known carbohydrate

linkages or the indirect detection of specific sugars, knowledge of the exact glycosyl composition opens the door to new detection possibilities.

The second main area that warrants further study is the Gram-positive / Gram-negative identity of *Pectinatus*. This organism represents a bridge between the classical split for bacteria. Proteins involved in the outer membrane of *Pectinatus* might contain clues as to how a “Gram-positive organism”, by genetics, developed an outer membrane. Work is currently beginning in the laboratory to sequence and characterize the major outer membrane protein found in *Pectinatus*. Will this protein be similar to S-layer proteins found in Gram-positive organisms or will it be closer to membrane-bound Gram-negative proteins? Can this protein reveal anything about the distinction between a Gram-positive and Gram-negative phenotype? Whatever the outcome, this study promises to shed more light on this unusual genus.

Finally, the third main project that is left to follow up on from this study is to develop a genus- and species-specific PCR detection protocol based on *Pectinatus* flagellin genes. In the discussion of the preliminary work already done, the two major tasks left have already been outlined. These are to expand the sample set used to design the PCR primers to include as many *Pectinatus* isolates as possible and expand the test set to include close relatives of *Pectinatus* that also spoil beer. Preliminary results indicate that specific detection of *Pectinatus* isolates in the brewery can be done with flagellin-specific PCR. A thorough study of this possible detection method is needed to bear this hypothesis out.

In the race to develop new methods to detect, identify and remove beer spoilage organisms from the brewery, taking a step back to understand the organisms targeted is sometimes forgotten. Detection methods used are only as good as the science behind

them. Taking the time to learn new characteristics about beer spoilage organisms, especially ones as understudied as *Pectinatus*, allows for new doors to be opened, new approaches to be considered and better detection strategies to be devised.

5.0 BIBLIOGRAPHY

- Altschul, S.F., W. Gish, W. Miller, E. W. Myers and D. J. Lipman. (1990) Basic local alignment search tool. *J. Mol. Biol.* 215:403-410.
- Amersham Biosciences. (2002) Handbook for GST Gene Fusion System, Chapter 3: Monitoring expression, optimizing growth and preparing large-scale cultures, Edition AA.
- Atsumi, T., L. McCarter and Y. Imae. (1992) Polar and lateral flagellar motors of marine *Vibrio* are driven by different ion-motive forces. *Nature.* 355:182-184.
- Barney M. and E. Kot. (1992) A comparison of rapid microbiological methods for detecting beer spoilage microorganisms. *MBAA Technical Quarterly.* 29(3):91-95.
- Brimer, C. D. and T. C. Montie. (1998) Cloning and comparison of *fliC* genes and identification of glycosylation in the flagellin of *Pseudomonas aeruginosa* a-type strains. *J. Bacteriol.* 180(12):3209-3217.
- Brinkman, F. S. L. and D. D. Leipe. (2001) Phylogenetic analysis. In *Bioinformatics: A practical guide to the analysis of genes and proteins.* A. D. Baxevanis and B. F. F. Ouellette (eds.). John Wiley & Sons, Inc., Toronto, ON, pp. 323-358.
- Day, A. (1987) Rapid detection of microbial spoilage. In *Brewing Microbiology.* F. G. Priest, and I. Campbell (eds.). 1st Edition. Elsevier Applied Science Publishers Ltd., New York, NY. pp. 207-231.
- Dayhoff, M. O., R. M. Schwartz and B. C. Orcutt. (1978) A model of evolutionary change in proteins. In *Atlas of protein sequence and structure.* M. O. Dayhoff (ed). Washington, D.C.: National Biomedical Research Foundation. pp. 345-362.
- De Man, J. C., M. Rogosa and M. E. Sharpe. (1960) A medium for the cultivation of lactobacilli. *J. Appl. Bacteriol.* 23:130-135.
- Dobson, C. M. (2001) Phylogenetic analysis of *Lactobacillus* and *Pediococcus* based on the 16S rRNA gene, 16S-23S rRNA interspacer region and heat shock 60 protein. M.Sc. Thesis, University of Saskatchewan.

- Dobson, C. M., H., Deneer, S. Lee, S. Hemmingsen, S. Glares and B. Ziola. (2002) Phylogenetic analysis of the genus *Pediococcus*, including *Pediococcus clausenii* sp. nov., a new lactic acid bacterium isolated from beer. Int. J. Syst. Evol. Microbiol. 52:2003-2010.
- Doetsh, R. N. and R. D. Sjoblad. (1980) Flagellar structure and function in eubacteria. Ann. Rev. Microbiol. 34:69-108.
- Dubray, G. and G. Bezar. (1982) A highly sensitive periodic acid-silver stain for 1,2-diol groups of glycoproteins and polysaccharides in polyacrylamide gels. Analytical Biochem. 119:325-329.
- Felsenstein, J. (1989) PHYLIP – Phylogeny Inference Package (Version 3.2). Cladistics. 5:164-166.
- Flahaut, S., Y. Tierny, D. Watier, J. P. Hornez and J. Jeanfils. (2000) Impact of thermal variations on biochemical and physiological traits in *Pectinatus* sp. Int. J. Food Microbiol. 55:53-61.
- Foster, A. and L. B. Andersen. (1999) Gas chromatographic technique for the confirmation of *Megasphaera* and *Pectinatus* spp. J. Am. Soc. Brew. Chem. 57(3):91-93.
- Fukunaga, M. and Y. Koreki. (1996) A phylogenetic analysis of *Borrelia burgdorferi* sensu lato isolates associated with Lyme disease in Japan by flagellin gene sequence determination. Int. J. Sys. Bacteriol. 46(2):416-421.
- Gares, S. L., M. S. Whiting, W. M. Ingeldew, and B. Ziola. (1993) Detection and identification of *Pectinatus cerevisiiphilus* using surface-reactive monoclonal antibodies in a membrane filter-based fluoroimmunoassay. J. Am. Soc. Brew. Chem. 51(4):158-163.
- Garrity, G. M. and J. G. Holt. (2001) The road map to the *manual*. In *Bergey's manual of systematic bacteriology*. Volume 1. D. R. Boone, R. W. Castenholz and G. M. Garrity (eds.). Springer Publishing, New York, N.Y. pp. 119-166.
- Geis, G., S. Suerbaum, B. Forsthoff, H. Lying and W. Opferkuch. (1993) Ultrastructure and biochemical studies of the flagellar sheath of *Helicobacter pylori*. J. Med. Microbiol. 38(5):371-377.
- Gilman, M. Z., J. L. Wiggs and M. J. Chamberlin. (1981) Nucleotide sequences of two *Bacillus subtilis* promoters used by *Bacillus subtilis* sigma-28 RNA polymerase. Nucleic Acids Res. 9(22):5991-6000.
- Graham, E. R. B. (1972) Fetuin. In *Glycoproteins, Their composition, structure and function. Part A*. A. Gottschalk (ed.). Elsevier Publishing Co., New York, NY. pp. 718-727.

- Guerry, P., S. M. Logan, S. Thornton and T. J. Trust. (1990) Genomic organization and expression of *Campylobacter* flagellin genes. J. Bacteriol. 172(4):1853-1860.
- Haikara, A. (1985) Detection of anaerobic, Gram-negative bacteria in beer. Monatsschrift für Brauwissenschaft. Heft 6:239-243.
- Haikara, A. (1992) The Genera *Pectinatus* and *Megasphaera*. In *The Prokaryotes: A handbook on the biology of bacteria: ecophysiology, isolation, identification, applications*. 2nd Edition. A. Balows, H. Truper, M. Dworkin, W. Harder, K. Schliefer (eds.). Springer-Verlag, New York, NY. pp. 1993-2004.
- Haikara, A., L. Penttilä, T. M. Enari and K. Lounatmaa. (1981a) Microbiological, biochemical, and electron microscopic characterization of a *Pectinatus* strain. Appl. Environ. Microbiol. 41(2):511-517.
- Haikara, A., T. M. Enari and K. Lounatmaa. (1981b) The genus *Pectinatus*, a new group of anaerobic beer spoilage bacteria. EBC Congress. 24:229-240.
- Hakalehto, E., H. Santa, J. Vepsäläinen, R. Laatikainen and J. Finne. (1997) Identification of a common structural motif in the disordered N-terminal region of bacterial flagellins. Eur. J. Biochem. 250:19-29.
- Hales, B. A., J. A. Morgan, C. A. Hart and C. Winstanley. (1998) Variation in flagellin genes and proteins of *Burkholderia cepacia*. J. Bacteriol. 180(5):1110-1118.
- Hansen, J. E., O. Lund, N. Tolstrup, A. A. Gooley, K. L. Williams and S. Brunak. (1998) NetOglyc: Prediction of mucin type O-glycosylation sites based on sequence context and surface accessibility. Glycoconjugate J. 15:115-130.
- Helander, I. M. and A. Haikara. (1995) Cellular fatty acyl and alkenyl residues in *Megasphaera* and *Pectinatus* species: contrasting profiles and detection of beer spoilage. Microbiol. 141:1131-1137.
- Helander, I., E. Hakalehto, J. Ahvenainen and A. Haikara. (1983) Characterization of lipopolysaccharides of *Pectinatus cerevisiiphilus*. FEMS Microbiol. Lett. 18:223-226.
- Helander, I. M., H. Moll and U. Zahringer. (1993) 4-O-(2-amino-2-deoxy- α -D-glucopyranosyl)-3-deoxy-D-manno-2-octulosonic acid, a constituent of lipopolysaccharides of the genus *Pectinatus*. Eur. J. Biochem. 213:377-381.
- Helander, I. M., I. Kilpeläinen, M. Varra, A. P. Moran, B. Lindner and U. Seydel. (1994) Chemical structure of the lipid A component of lipopolysaccharides of the genus *Pectinatus*. Eur. J. Biochem. 224:63-70.

- Helander, I. M., R. Hurme, A. Haikara and A. P. Moran. (1992) Separation and characterization of two chemically distinct lipopolysaccharides in two *Pectinatus* species. *J. Bacteriol.* 174(10):3348-3354.
- Helmann, J. D. and M. J. Chamberlin. (1987) DNA sequence analysis suggests that expression of flagellar and chemotaxis genes in *Escherichia coli* and *Salmonella typhimurium* is controlled by an alternative σ factor. *Proc. Natl. Acad. Sci. USA* 84:6422-6424.
- Homma, M., H. Fujita, S. Yamaguchi and T. Iino. (1987) Regions of *Salmonella typhimurium* flagellin essential for its polymerization and excretion. *J. Bacteriol.* 169(1):291-296.
- Hounsell, E. F., M. J. Davies and K. D. Smith. (1996) Chemical methods of analysis of glycoproteins. In *The Protein Protocols Handbook*. J. M. Walker (ed.). Humana Press Inc., Totowa, NJ. pp. 633-634.
- Ide, N., T. Ikebe and K. Kutsukake. (1999) Reevaluation of the promoter structure of the class 3 flagellar operons of *Escherichia coli* and *Salmonella*. *Genes Genet. Syst.* 74:113-116.
- Juvonen, R., R. Satokari, K. Mallison and A. Haikara. (1999) Detection of spoilage bacteria in beer by polymerase chain reaction. *J. Am. Soc. Brew. Chem.* 57(3):99-103.
- Kelly-Wintenberg, K., S. L. South and T. C. Montie. (1993) Tyrosine phosphate in a- and b-type flagellins of *Pseudomonas aeruginosa*. *J. Bacteriol.* 175(8):2458-2461.
- Kimura, M. (1980) A simple method for estimating evolutionary rate of base substitutions through comparative studies of nucleotide sequences. *J. Molecular Evolution* 16:111-120.
- Kuwajima, G. (1988) Construction of a minimum-size functional flagellin of *Escherichia coli*. *J. Bacteriol.* 170(7):3305-3309.
- Kuwajima, G., J. I. Asaka, T. Fujiwara, T. Fujiwara, K. Node and E. Kondo. (1986) Nucleotide sequence of the *hag* gene encoding flagellin of *Escherichia coli*. *J. Bacteriol.* 168(3):1479-1483.
- Laemmli, U. K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* 227:680-685.
- Lechner, J. and F. Wieland. (1989) Structure and biosynthesis of prokaryotic glycoproteins. *Annu. Rev. Biochem.* 58:173-194.
- Lee, S. Y. (1994) SMMP – A medium for selective isolation of *Megasphaera* and *Pectinatus* from the brewery. *J. Am. Soc. Brew. Chem.* 52(3):115-119.

- Lee, S. Y., M. S. Mabee and N. O. Jangaard. (1978) *Pectinatus*, a new genus of the family *Bacteroidaceae*. *Int. J. Syst. Bacteriol.* 28(4):582-594.
- Lee, S. Y., M. S. Mabee, N. O. Jangaard and E. K. Horiuchi. (1980) *Pectinatus*, a new genus of bacteria capable of growth in hopped beer. *J. Inst. Brew.* 86:28-30.
- Lee, S. Y., S. E. Moore and M. S. Mabee. (1981) Selective-differential medium for isolation and differentiation of *Pectinatus* from other brewery microorganisms. *Appl. Environ. Microbiol.* 41(2):386-387.
- Limberger, R. J., L. L. Slivienski, D. B. Yelton and N. W. Charon. (1992) Molecular genetic analysis of a class B periplasmic flagellin gene of *Treponema phagedenis*. *J. Bacteriol.* 174(20):6404-6410.
- Losick, R. and J. Pero. (1976) Regulatory subunits of RNA polymerase. In *RNA Polymerase*. Losick, R. and M. Chamberlin (eds.). Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, NY, pp. 227-246.
- Lovgren, A., M. Y. Zhang, A. Engstrom and R. Landen. (1993) Identification of two expressed flagellin genes in the insect pathogen *Bacillus thuringiensis* subsp. *alesti*. *J. Gen. Microbiol.* 139:21-30.
- Makrides, S. C. (1996) Strategies for achieving high-level expression of genes in *Escherichia coli*. *Microbiol. Rev.* 60(3):512-538.
- Marchler-Bauer, A., J. B. Anderson, C. DeWeese-Scott, N. D. Fedorova, L. Y. Geer, S. He, D. I. Hurwitz, J. D. Jackson, A. R. Jacobs, C. J. Lanczycki, C. A. Liebert, C. Liu, T. Madej, G. H. Marchler, R. Mazumder, A. N. Nikolskaya, A. R. Panchenko, B. S. Rao, B. A. Shoemaker, V. Simonyan, J. S. Song, P. A. Thiessen, S. Vasudevan, Y. Wang, R. A. Yamashita, J. J. Yin and S. H. Bryant. (2003) CDD: a curated Entrez database of conserved domain alignments. *Nucleic Acids Res.* 31(1):383-387.
- Martin, J. H. and D. C. Savage. (1988) Cloning, nucleotide sequence, and taxonomic implications of the flagellin gene of *Roseburia cecicola*. *J. Bacteriol.* 170(6):2612-2617.
- McCarter, L. L. (1995) Genetic and molecular characterization of the polar flagellum of *Vibrio parahaemolyticus*. *J. Bacteriol.* 177(6):1595-1609.
- McGee, K., P. Horstedt and D. L. Milton. (1996) Identification and characterization of additional flagellin genes from *Vibro anguillarum*. *J. Bacteriol.* 178(9):5188-5198.
- Millikan, D. S., H. Felbeck and J. L. Stein. (1999) Identification and characterization of a flagellin gene from the endosymbiont of the hydrothermal vent tubeworm *Riftia pachyptila*. *Appl. Environ. Microbiol.* 65(7):3129-3133.

- Mimori-Kiyosue, Y., F. Vonderviszt, I. Yamashita, Y. Fujiyoshi and K. Namba. (1996) Direct interaction of flagellin termini essential for polymorphic ability of flagellar filament. *Proc. Natl. Acad. Sci. USA.* 93:15108-15113.
- Mirel, D. B. and M. J. Chamberlin. (1989) The *Bacillus subtilis* flagellin gene (*hag*) is transcribed by the σ^{28} form of RNA polymerase. *J. Bacteriol.* 171(6):3095-3101.
- Montgomery, R. (1970) Glycoproteins; Glycoproteins having one carbohydrate group in each protein unit; Ovalbumin. In *The Carbohydrates; Chemistry and Biochemistry*. W. Pigman and D. Horton (eds.) Academic Press, New York, N. Y. pp. 627-709.
- Moore, K. L. (2003) The biology and enzymology of protein tyrosine O-sulfation. *J. Biological Chem.* 278(27):24243-24246.
- Motoyama, Y. and T. Ogata. (2000a) 16S-23S rDNA spacer of *Pectinatus*, *Selenomonas* and *Zymophilus* reveal new phylogenetic relationships between these genera. *Int. J. Syst. Evol. Microbiol.* 50:883-886.
- Motoyama, Y. and T. Ogata. (2000b) Detection of *Pectinatus* spp. by PCR using 16S-23S rDNA spacer regions. *J. Am. Soc. Brew. Chem.* 58(1):4-7.
- Munroe, D. J., M. Haas, E. Bric, T. Whitton, H. Aburatani, K. Hunter, D. Ward and D. E. Housman. (1994) IRE-Bubble PCR: A rapid method for efficient and representative amplification of human genomic DNA sequences from complex sources. *Genomics.* 19:506-514.
- Nicholas, K. B., H. B. Nicholas Jr. and D. W. Deerfield II. (1997) GeneDoc: Analysis and Visualization of Genetic Variation. *EMBNEW.NEWS* 4:14.
- Nuijten, P. J., F. J. Van Asten, W. Gaastra and B. A. Van der Zeijst. (1990) Structural and functional analysis of two *Campylobacter jejuni* flagellin genes. *J. Biol. Chem.* 265(29):17798-17804.
- Olsen, G. J. and C. R. Woese. (1993) Ribosomal RNA: a key to phylogeny. *FASEB J.* 7:113-123.
- Panelius, J., P. Lahdenne, H. Saxen, T. Heikkila and I. Seppala. (2001) Recombinant flagellin A proteins from *Borrelia burgdorferi* sensu stricto, *B. afzelii*, and *B. garinii* in serodiagnosis of Lyme Borreliosis. *J. Clin. Microbiol.* 39(11):4013-4019.
- Priest, F. G. (1996) Anaerobic Gram-negative rods. In *Brewing Microbiology*. F. G. Priest, and I. Campbell (eds.). 2nd Edition. Elsevier Applied Science Publishers Ltd., New York, NY. pp. 180-183.
- Racusen, D. (1979) Glycoprotein detection in polyacrylamide gel with thymol and sulfuric acid. *Analytical Biochem.* 99:474-476.

- Riley, J., R. Butler, D. Ogilvie, R. Finniear, D. Jenner, S. Powell, R. Anand, J. C. Smith and A. F. Markham. (1990) A novel, rapid method for the isolation of terminal sequences from yeast artificial chromosome (YAC) clones. *Nucleic Acids Res.* 18(10):2887-2890.
- Rozen, S. and H. J. Skaletsky. (2000) Primer3 on the WWW for general users and for biologist programmers. In: *Bioinformatics Methods and Protocols: Methods in Molecular Biology*. S. Krawetz and S. Misener (eds.). Humana Press, Totowa, NJ, pp. 365-386.
- Satokari, R., R. Juvonen, A. Von Wright, and A. Haikara. (1997) Detection of *Pectinatus* beer spoilage bacteria by using the polymerase chain reaction. *J. Food Protection*. 60(12):1571-1573.
- Satokari, R., R. Juvonen, K. Mallison, A. Von Wright and A. Haikara. (1998) Detection of beer spoilage bacteria *Megasphaera* and *Pectinatus* by polymerase chain reaction and colorimetric microplate hybridization. *Int. J. Food. Microbiol.* 45:119-127.
- Scharf, B., H. Schster-Wolff-Buhring, R. Rachel and R. Schmitt. (2001) Mutational analysis of the *Rhizobium lupine* H13-3 and *Sinorhizobium meliloti* flagellin genes: importance of flagellin A for flagellar filament structure and transcription regulation. *J. Bacteriol.* 183(18):5334-5342.
- Schleifer, K. H., M. Leuteritz, N. Weiss, W. Ludwig, G. Kirchhof and H. Seidel-Rufer. (1990) Taxonomic study of anaerobic, Gram-negative, rod-shaped bacteria from breweries: Emended description of *Pectinatus cerevisiiphilus* and description of *Pectinatus frisingensis* sp. nov., *Selenomonas lactericifex* sp. nov., *Zymophilus raffinovorans* gen. nov., sp. nov., and *Zymophilus paucivorans* sp. nov. *Int. J. Syst. Bacteriol.* 40(1):19-27.
- Senchenkova, S. N., A. S. Shashkov, A. P. Moran, I. M. Helander and Y. A. Knirel. (1995) Structures of the O-specific polysaccharide chains of *Pectinatus cerevisiiphilus* and *Pectinatus frisingensis* lipopolysaccharides. *Eur. J. Biochem.* 232:552-557.
- Shah, D. S. H., T. Perehinec, S. M. Stevens, S. I. Aizawa and R. E. Sockett. (2000) The flagellar filament of *Rhodobacter sphaeroides*: pH-induced polymorphic transitions and analysis of the *fliC* gene. *J. Bacteriol.* 182(18):5218-5224.
- Snyder, L. and W. Champness. (1997) In *Molecular Genetics of Bacteria*. Chapter 2: Introduction to macromolecular synthesis: Gene expression. ASM Press, Washington, D. C., pp. 43-72.
- Spangenberg, C., T. Heuer, C. Burger and B. Tummeler. (1996) Genetic diversity of flagellins of *Pseudomonas aeruginosa*. *FEMS Lett.* 396:213-217.

- Stackebrandt, E., H. Pohla, R. Kroppensedt, H. Hippe and C. R. Woese. (1985) 16S rRNA analysis of *Sporomusa*, *Selenomonas* and *Megasphaera*: on the phylogenetic origin of gram-positive eubacteria. Arch. Microbiol. 143:270-276.
- Suihko, M. L. and A. Haikara. (2001) Characterization of *Pectinatus* and *Megasphaera* strains by automated ribotyping. J. Int. Brew. 107(3):175-184.
- Tasteyre, A., M. C. Barc, T. Karjalainen, P. Dodson, S. Hyde, P. Bourlioux and P. Borriello. (2000) A *Clostridium difficile* gene encoding flagellin. Microbiol. 146:957-966.
- Tatusov R. L., D. A. Natale, I. V. Garkavtsev, T. A. Tatusova, U. T. Shankavaram, B. S. Rao, B. Kiryutin, M. Y. Galperin, N. D. Fedorova and E. V. Koonin. (2001) The COG database: new developments in phylogenetic classification of proteins from complete genomes. Nucleic Acids Res. 29(1):22-8.
- Thibault, P., S. M. Logan, J. F. Kelly, J. R. Brisson, C. P. Ewing, T. J. Trust and P. Guerry. (2001) Identification of the carbohydrate moieties and glycosylation motifs in *Campylobacter jejuni* flagellin. J. Biol. Chem. 276(37):34862-34870.
- Tholozan, J. L., J. M. Membre and J. P. Grivet. (1997) Physiology and development of *Pectinatus cerevisiiphilus* and *Pectinatus frisingensis*, two strict anaerobic beer spoilage bacteria. Int. J. Food Microbiol. 35:29-39.
- Thomas, N. A., S. L. Bardy and K. F. Jarrell. (2001) The archaeal flagellum: a different kind of prokaryotic motility structure. FEMS Microbiol. Rev. 25:147-174.
- Thompson, J. D., D. G. Higgins and T. J. Gibson. (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position specific gap penalties and weight matrix choice. Nucleic Acids Res. 22:4673-80.
- Thompson, J. D., T. J. Gibson, F. Plewniak, F. Jeanmougin and D. G. Higgins. (1997) The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res. 24:4876-4882.
- Uchida, Y., Y., Tsukada and T. Sigimori. (1979) Enzymatic properties of neuraminidase from *Arthrobacter ureafaciens*. J. Biochem. 86(5):1573-1585.
- Wallich, R., S. E. Moter, M. M. Simon, K. Ebnet, A. Heiberger and M. D. Kramer. (1990) The *Borrelia burgdorferi* flagellum-associated 41-kilodalton antigen (flagellin): molecular cloning, expression, and amplification of the gene. Infection and Immunity. 58(6):1711-1719.
- Wieland, F., G. Paul and M. Sumper. (1985) Halobacterial flagellins are sulfated glycoproteins. J. Biol. Chem. 260(28):15180-15185.

- Willems, A. and M. D. Collins. (1995) Phylogenetic placement of *Dialister pneumosintes* (formerly *Bacteroides pneumosintes*) within the *Sporomusa* subbranch of the *Clostridium* subphylum of the Gram-positive bacteria. *Int. J. Syst. Bacteriol.* 45(2):403-405.
- Wilson, D. R. and T. J. Beveridge. (1993) Bacterial flagellar filaments and their component flagellins. *Can. J. Microbiol.* 39:451-472.
- Winstanley, C. and J. A. W. Morgan. (1997) The bacterial flagellin gene as a biomarker for detection, population genetics and epidemiological analysis. *Microbiol.* 143:3071-3084.
- Yarnell W. S. and J. W. Roberts. (1999) Mechanism of intrinsic transcription termination and antitermination. *Science.* 284:611-615.
- Yasuhara, T., T. Yuuki and N. Kagami. (2001) Novel quantitative method for detection of *Pectinatus* using rRNA targeted fluorescent probes. *J. Am. Soc. Brew. Chem.* 59(3):117-121.

6.0 APPENDIX A - Solutions, Buffers and Media

Acetate buffer, 0.1 M, pH 5.5: 3.54 g anhydrous sodium acetate and 393 μ l glacial acetic acid in 500 ml rdH₂O.

Antibody buffer: PBS, 2% (w/v) BSA, 2% (v/v) Tween 20 and 0.2 mM merthiolate.

Bacterial plasmid mini-prep solution: 8% sucrose, 0.5% (v/v) Triron-x-100, 50 mM EDTA (pH 8.0) and 10 mM Tris-HCl (pH 8.0).

Blocking solution: PBS, 3% (w/v) BSA and 0.5 mM merthiolate.

Coomassie Blue destain: 100 ml glacial acetic acid, 400 ml methanol and 1.5 L rdH₂O.

Coomassie Blue stain: 500 ml 95% ethanol, 100 ml glacial acetic acid, 500 ml rdH₂O and 1.1 g coomassie blue R250. Filter as needed.

Luria-Bertani (LB) + ampicillin:

5.0 g tryptone, 2.5 g yeast extract and 5.0 g sodium chloride in 500 ml rdH₂O; for plates, add 7.5 g agar; autoclave and cool to 50⁰C before adding 1.0 ml of 25 mg/ml ampicillin (in rdH₂O).

Lysozyme: 10 mg/ml in 10 mM Tris-HCl, pH 8.0.

Mannitol medium: 0.5 g yeast extract, 0.3 g peptone and 2.5 g D-mannitol in 100 ml rdH₂O; for plates, add 1.5 g agar; autoclave.

MRS agar: 66.2 g MRS agar powder in rdH₂O to a final volume of 1.0 L; autoclave for 20 min at 121⁰C.

MRS media: 52.2 g MRS broth powder in rdH₂O to a final volume of 1.0 L; autoclave for 20 min at 121⁰C.

Nutrient medium: 2.5 g nutrient broth powder in 100 ml rdH₂O; for plates, add 1.5 g agar; autoclave for 20 min at 121⁰C.

PCR buffer, 10x: 200mM Tris, pH 8.4, 500 mM KCl, 15 mM MgCl₂ and 0.01% (w/v) gelatin.

Stock buffer without MgCl₂: Dissolve 1.21 g Tris and 1.87 g KCl₂; pH solution to 8.4 with HCl; add 5 mg of gelatin and heat gently until dissolved; bring volume to 35 ml with rdH₂O.

Working buffer: Combine 3.5 ml of stock buffer and 1.5 ml of 50 mM MgCl₂ solution (Invitrogen – supplied with Taq enzyme); sterilize with 0.22 µM filter; aliquot into 100 µl volumes; store at -20°C.

Phosphate-buffered saline (PBS) pH 7.2:

Stock PBS solution: 54.8 g disodium hydrogen orthophosphate and 15.8 g sodium dihydrogen orthophosphate, bring to 1.0 L with rdH₂O

Working solution: 40 ml stock PBS solution and 8.5 g sodium chloride, bring to 1.0 L with rdH₂O; autoclave 20 min at 121°C.

SDS-PAGE dissociation solution:

4:1:1 ratio of 5x stock; Bromphenol blue tracking dye; β-mercaptoethanol, respectively.

5x stock: 1.0 g SDS into 10 ml SDS-PAGE stacking gel buffer.

Bromphenol blue tracking dye: 5.0 mg bromphenol blue in 20 ml of 50% glycerol/50% rdH₂O.

SDS-PAGE gel reagents:

Stock acrylamide: 30.0 g acrylamide and 0.8 g N,N'-methylene-bis-acrylamide in 100 ml of rdH₂O.

EDTA, 0.2 M: 0.745 g EDTA into 10 ml rdH₂O.

Separating gel buffer: 18.17 g Tris into 70 ml rdH₂O, pH to 8.8 with HCl; add 0.4 g SDS; bring final volume to 100 ml.

Stacking gel buffer: 6.06 g Tris into 70 ml rdH₂O, pH to 6.8 with HCl; add 0.4 g SDS; bring final volume to 100 ml.

SDS-PAGE running buffer:

10x stock: 60 g Tris and 288 g glycine in 2.0 L rdH₂O.

1x working solution: dilute 10x stock 1:10 with rdH₂O and add 1.0 g SDS per liter of final volume.

SDS-PAGE separating gel (12%): 11.33 ml stock acrylamide, 6.82 ml separating gel buffer, 270 µl 0.2 M EDTA, 8.70 ml rdH₂O, 59.4 µl 10% (w/v) ammonium persulfate and 35.5 µl N,N,N',N'-tetramethylethylenediamine (TEMED).

SDS-PAGE stacking gel (4%): 860 µl stock acrylamide, 1.63 ml stacking gel buffer, 65 µl 0.2 M EDTA, 3.97 ml rdH₂O, 32.5 µl 10% (w/v) ammonium persulfate and 16.3 µl TEMED.

SOC medium: 31.5 g SOC medium powder into 1.0 L rdH₂O; autoclave for 20 min at 121°C.

SSC/Tris, 2x: 100 ml of 10 mM Tris-HCl, pH 8.0, containing 1.753 g sodium chloride and 0.882 g of sodium citrate.

TBE buffer, 10x: 108 g Tris, 55 g Boric Acid and 9.3 g EDTA, bring to 1.0 L with rdH₂O.

TBE buffer, 1x: 1:10 dilution of 10x TBE buffer with rdH₂O.

TBS: 20 mM Tris-HCl, pH 7.5, 500 mM NaCl.

TE buffer, pH 7.6: 10 mM Tris-HCl, pH 7.6, 1 mM EDTA, pH 8.0.

TE buffer, ¼, pH 7.6: 1:4 dilution of TE buffer, pH 7.6, with rdH₂O.

Tracking dye: 50% (v/v) glycerol, 50% (v/v) rdH₂O, containing 0.0025% Bromphenol Blue.

Transfer Buffer: 25 mM Tris, 192 mM glycine and 20% methanol.

TTBS: TBS containing 0.05% (v/v) Tween 20.

7.0 APPENDIX B

7.1 Chemical, media and reagents

100 bp DNA ladder	Invitrogen, Burlington, ON
2' Deoxyadenosine 5'-triphosphate	Amersham Biosciences, Baie d'Urfe, QB
2' Deoxycytosine 5'-triphosphate	Amersham Biosciences
2' Deoxyguanosine 5'-triphosphate	Amersham Biosciences
2' Deoxythymidine 5'-triphosphate	Amersham Biosciences
Acetic acid (glacial)	EMD Chemicals Inc., Gibbstown, NJ
Acrylamide	Bio-Rad Laboratories, Mississauga, ON
Agar	Difco, Becton Dickson, Sparks, MD
Agarose	Invitrogen
Alcian blue	Bio-Rad Laboratories
Alkaline phosphatase and buffer	Sigma Chemical Co., St. Louis, MO
Alkaline phosphatase-conjugated Affinipure TM goat anti-rabbit IgG (H+L) (Jackson ImmunoResearch Laboratories)	Cederlane Laboratories Ltd, Hornby, ON
Alkaline phosphatase-conjugated Anti-goat IgG H+L chain (rabbit) (Calbiochem)	EMD Chemicals Inc.
α -chymotrypsin (α -chymotrypsinogen-A from bovine pancreas)	Sigma Chemical Co.
α -galactosidase (from <i>Aspergillus niger</i>)	Sigma Chemical Co.
α -L-Fucosidase (from bovine kidney)	Sigma Chemical Co.

Ammonium persulphate	Bio-Rad Laboratories
Ampicillin	Invitrogen
BCA Protein Assay, Microwell plate protocol (Pierce)	Professional Diagnostics, Edmonton, AB
BCIP/NBT color development system	Bio-Rad Laboratories
β -galactosidase (from Jack Beans)	Sigma Chemical Co.
β -mercaptoethanol	Sigma Chemical Co.
β -N-acetylglucosaminidase (from <i>Aspergillus niger</i>)	Sigma Chemical Co.
BLOT-FastStain	Chemicon International, Temecula, CA
Boric acid	EMD Chemicals Inc.
Bovine serum albumin (BSA)	Sigma Chemical Co.
Bromphenol blue	Bio-Rad Laboratories
Chloroform	VWR Canlab, Mississauga, ON
Coomassie blue R250	Sigma Chemical Co.
DH10 β <i>E. coli</i> cells	Dr. Wei Xiao laboratory
Disodium hydrogen orthophosphate	VWR Canlab
Ethanol, 95%	VWR Canlab
Ethidium bromide	EMD Chemicals Inc.
Ethylenediaminetetraacetic acid (EDTA)	Invitrogen
Fetuin	Sigma Chemical Co.
Gelatin	Bio-Rad Laboratories
Glutathione S-transferase (GST) Gene Fusion System	Amersham Biosciences
Glycerol	VWR Canlab

Glycine	EMD Chemicals Inc.
Horse radish peroxidase (HRP)	Sigma Chemical Co.
Hydrochloric acid	VWR Canlab
Isopropylthio-B-D-galactoside (IPTG)	Invitrogen
Isoamyl alcohol	Fisher Scientific, Ottawa, ON
Isopropyl alcohol	EMD Chemicals Inc.
Low DNA mass ladder	Invitrogen
Lysozyme	Sigma Chemical Co.
Magnesium chloride, 50 mM solution	Invitrogen
Mannitol (-D)	Difco
Methanol	EMD Chemicals Inc.
MRS medium/agar	VWR Canlab
N, N'-methylene-bis-arcylamide	Bio-Rad Laboratories
N, N, N', N'-tetramethylethylenediamine (TEMED)	Bio-Rad Laboratories
Neuraminidase (from <i>Vibrio cholerae</i>) (Calbiochem)	EMD Chemicals Inc.
Nutrient broth	Oxoid, Basingstoke, Hampshire, England
Ovalbumin	Worthington Biochemical, Freehold, NJ
PCR primers	Sigma Chemical Co.
Peptone	Difco
Periodic acid	Sigma Chemical Co.
Phenol	Invitrogen
Platinum Taq DNA Polymerase	Invitrogen

PNGase F and accompanying enzyme reagents	New England Biolabs, Mississauga, ON
Prestained protein molecular weight standards, high range	Invitrogen
Potassium chloride	Fisher Scientific
Puregene Genomic DNA purification kit (Gentra Systems)	Inter-Medico, Markham, ON
Restriction enzymes: BamHI, EcoRV, HaeIII, RsaI, SalI, Sau3AI and accompanying buffers and reagents	New England Biolabs
Schiff's reagent	Sigma Chemical Co.
SOC medium	American Biorganics, Niagara Falls, NY
Sodium chloride	VWR Canlab
Sodium citrate	Baker Chemicals, Phillipsburg NJ
Sodium dihydrogen orthophosphate	VWR Canlab
Sodium dodecyl sulphate (SDS)	VWR Canlab
Sodium periodate	VWR Canlab
Sucrose	VWR Canlab
Sulfuric acid	EMD Chemicals Inc.
T4 DNA ligase (and reaction buffer)	Invitrogen
Thymol	Sigma Chemical Co.
Tris	VWR Canlab
Triton-x-100	Sigma Chemical Co.
Tryptone	Difco
Tween 20	ICN Biomedicals, Aurora, OH
Wizard PCR Preps DNA Purification System (Promega)	Invitrogen

Yeast extract

Invitrogen

7.2 Equipment and supplies

Beckman L8-70M ultracentrifuge

Beckman Coulter Canada, Mississauga, ON

Blotting paper

Schleich and Schuell, Keene, NH

C10 platform shaker

New Brunswick Scientific, Edison, NJ

Cheesecloth

No Name, Superstore, Saskatoon, SK

E. coli Pulser

Bio-Rad Laboratories

E. coli Pulser Transformation apparatus

Bio-Rad Laboratories

EIA reader, model 2550

Bio-Rad Laboratories

Eppendorf 5415C table-top centrifuge

Brinkmann Instruments, Mississauga, ON

Filters, 0.22 µm
(Pall Gelman Laboratory)

VWR Canlab

Gas-pak jars

BD Biosciences, Mississauga, ON

Gel doc 1000 system

Bio-Rad Laboratories

Incubator, model 322

NARCO / Tri-Star, Cincinnati, OH

Mini-subTM DNA cell agarose gel tank

Bio-Rad Laboratories

PCR Sprint Temperature Cycling
System

Hybaid, Franklin, MA

Pipette tips (2-200 µl and 100-1000 µl)

DiaMed, Mississauga, ON

Pipettors (all sizes)

Brinkmann Instruments

Plugged pipette tips (0.1-2.0 µl)

DiaMed

Plugged pipette tips (2-100 µl; 100-1000 µl)

Sarstedt Inc., St-Léonard, QB

Polyvinylidene difluoride (PVDF) membrane	Bio-Rad Laboratories
Power supply 1000/500 (for SDS-PAGE gels)	Bio-Rad Laboratories
Power supply, model 250 (for agarose gels)	Invitrogen
Protean II xi cell SDS-apparatus	Bio-Rad Laboratories
PTC-100 Programmable Thermal Controller	MJ Research, Waterton, MA
Rocker	Reliable Scientific, Inc., Hernando, MS
Sorvall Omnimixer	Du Pont, Mississauga, ON
Sorvall RC-5B refrigerated superspeed centrifuge (SA 600 rotor)	Du Pont
Transiluminator	Ultra-Violet Products, San Gabriel, CA
Tubes, 0.5 ml thick-walled	Rose Scientific Ltd., Edmonton, AB
Tubes, 0.5 ml thin-walled	Gordon Technologies, Mississauga, ON
Tubes, 1.5 ml	VWR Canlab
Versa-bath water bath	Fisher Scientific
Wide mini-sub TM DNA cell agarose gel tank	Bio-Rad Laboratories

7.3 Software used

BLAST¹; Altschul *et al.*, 1990; <http://www.ncbi.nlm.nih.gov/BLAST/>

CDD¹ (conserved domain database); Marchler-Bauer *et al.*, 2003;
<http://www.ncbi.nlm.nih.gov/Structure/cdd/cdd.shtml>

ClustalX²; Thompson *et al.*, 1994/1997; <http://inn-prot.weizmann.ac.il/software/>

DNA calculator¹; Unpublished; <http://www.sigma-genosys.com/>

EMBOSS¹; Unpublished; <http://bioinfo.pbi.nrc.ca:8090/EMBOSS/>

GeneDoc²; Nicholas *et al.*, 1997; <http://www.psc.edu/biomed/genedoc/>

MacReader II; Bio-Rad Laboratories.

Multi Analyst Software, version 1.1, build 34; Bio-Rad Laboratories.

PHYLP²; Felsenstein *et al.*, 1989; <http://evolution.genetics.washington.edu/phylip.html>

Primer3¹; Rozen and Skaletsky, 1998; http://cbr-rbc.nrc-cnrc.gc.ca/cgi-bin/primer3_www.cgi

TreeExplorer²; Unpublished; http://evolgen.biol.metro-u.ac.jp/TE/TE_man.html

¹ Software available for use on-line.

² Software available for free download to personal computer.